DOCUMENT RESUME

ED 155 059	SE 024 351

AUTHOR	Crocker, R. K.; And Others
TITLE	Teacher Control, Classroom Behaviors and Pupil Performance in Elementary Science Classes.
INSTITUTION	Memorial Univ., St. John's (Newfoundland).
PUB DATE	77
NOTE	311p.; Contains occasional light and broken type
EDRS PRICE	MF-$0.83 HC-$16.73 Plus Postage.
DESCRIPTORS	*Classroom Environment; *Classroom Research; *Educational Research; Elementary Education; *Elementary School Science; *Science Activities; Science Education; Teacher Role; Teaching Styles
IDENTIFIERS	Research Reports

ABSTRACT

This document reports on an educational research study addressing problems of classroom dynamics in the elementary level science laboratory. Major aims included identification of student and teacher roles, role departures, and pupil thought processes through pupil behaviors. Results indicate that changes in teacher and pupil roles do not necessarily follow even from drastic changes in setting and that role shifts may occur over time and can ultimately influence outcomes. (SL)

****
*   Reproductions supplied by EDRS are the best that can be made   *
*   from the original document.   *
****

ED155059

TEACHER CONTROL, CLASSROOM BEHAVIORS AND PUPIL PERFORMANCE IN ELEMENTARY SCIENCE CLASSES

R.K. Crocker, R.P. Amaria, H. Banfield, G.W. Clark,
W.F. Oakley, G. Rumboldt, D.B. Sheppard

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Robert K. Crocker

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

IERD

INSTITUTE FOR EDUCATIONAL
RESEARCH AND DEVELOPMENT
MEMORIAL UNIVERSITY OF NEWFOUNDLAND

SE 024 351

TABLE OF CONTENTS

Page

LIST OF TABLES i

ABSTRACT iv

ACKNOWLEDGEMENTS viii

Chapter

I. THE PROBLEM 1

  Objectives of the Study 1

  Rationale 2

  Scientific Processes 7

  Overview of Method 10

II. RELATED RESEARCH 12

  Teacher Control 12

  Small Group Interaction in the Classroom 18

  Pupil Characteristics, Classroom Behaviors, and Outcomes 20

III. METHOD 25

  Samples and School Settings 25

  The Curriculum 27

  Data Collection - Descriptive Study 28

  Experimental Design 30

  Data Sources - Experimental Study 34

  Pupil Classroom Behavior 36

  Data Analysis 39

Chapter Page

IV. INSTRUMENTS 43

Pupil Background Variables 43

Pupil Ability 44

Personality and Perceptions 46

Science Process Achievement 48

Pupil Perceptions and Preferences 49

Classroom Observation System 52

V. TEACHER AND PUPIL BEHAVIORS: AN OVERVIEW 69

Teacher Behaviors 71

Pupil Behaviors and Pupil-Pupil Interactions . . . 84

Teacher Interventions 102

Summary and Discussion 118

VI. EXPERIMENTAL RESULTS 121

Implementation of Treatments 123

Achievement. 126

Preference Scale 129

Choice Task 131

Interview Schedule 136

Pupil Classroom Behaviors 141

Chapter Page

VII. SUMMARY AND CONCLUSIONS 150

Summary of Purposes and Methodology 150

Summary of Descriptive Phase 152

Summary of Experimental Findings 158

Conclusions 162

A Behavior Setting Interpretation 172

Remarks on Methodology 175

FOOTNOTES 180

APPENDIX A - INSTRUMENTS 182

APPENDIX B - CODING MANUAL, CLASSROOM OBSERVATION SYSTEM . . 241

APPENDIX C - SUMMARIES OF REGRESSION ANALYSES FOR CLASSROOM BEHAVIORS: EXPERIMENTAL STUDY . . 281

REFERENCES 292

LIST OF TABLES

Table Page

1. Summary of Variables Used in Differentiating Experimental Treatments 31

2. Sample Comparison Matrix: Pedagogical Moves 66

3. Sample Percentages of Agreement Between Coder and Reviewer 68

4. Some Preliminary Indicators of Teacher Control of Lessons 73

5. Means and Ranges for Proportions of Each Pedagogical Move in Teacher Talk 75

6. Comparison of Proportions of Total Teacher Talk for Each Pedagogical Move for Four Studies . . 76

7. Proportions of Declarative, Interrogative and Imperative Sentences in Teacher Talk. 78

8. Proportions of Rating Reactions by Teachers 79

9. Proportion of Controlling Statements by Teachers 81

10. Proportions of Type of Reference Statements by Teachers 82

11. Proportions of Teacher Sentences in Substantive-Logical Categories by Grade 85

12. Proportions of Sentences Uttered by Individual Pupils by Receiver and Grade 88

13. Proportions of Each Pedagogical Move for Pupils 91

14. Proportions of Pupil Sentences in Substantive-Logical Categories by Grade 94

15. Proportions of Declarative, Interrogative and Imperative Sentences for Pupils 95

Table Page

16. Proportions of Pupil Sentences in Controlling Categories 97

17. Proportions of Type of Reference Statements for Pupils by Grade 99

18. Proportions of Rating Reactions for Pupils by Grade 100

19. Summary of Findings for Intervention Length and Number 105

20. Proportions of Teacher and Pupil Talk During Interventions and During Teacher-Class Interactions 110

21. Mean Proportions of Teacher and Pupil Discourse for Each Pedagogical Move 112

22. Proportions of Statements, Requests, and Commands for Teachers and Pupils 114

23. Proportions of Type of Reference Categories for Teachers and Pupils 116

24. Proportions of Ratings for Teachers and Pupils . . . 117

25. Comparison of Behaviors in High and Low Control Treatments 124

26. Treatment by Teacher Comparisons for Selected Teacher Behaviors 127

27. Summary of Regression Analysis, Science Process Achievement 128

28. Summary of Regression Analysis, Preference Scale . . 130

29. Cell Means for Treatment-Teacher Interactions, Preference Scale 132

30. Frequencies of Pupil Choice of High and Low Control Tasks 133

Table Page

31. Summary of Chi-Square Tests for Pupil Choice of High or Low Control Task by Pupil Characteristics 135

32. Summary of Pupil Responses to Interview Item Clusters Based on Four Rounds 137

33. Summary of Chi-Square Tests for Pupil Responses to Interview Items Cluster One for Rounds One to Four 139

34. Summary of Regression Analysis, Treatment x Teacher, Pupil Behavior Variables 145

35. Correlations Between Pupil Behavior Variables . . . 147

36. Cell Means, Treatment x Teacher, Pupil Behavior Variables 148

ABSTRACT

This report describes a two phase study designed to address a number of problems of classroom dynamics in the relatively unconventional setting of science laboratory activities in elementary school classrooms. The classes of interest were characterized by small group activities, with each group having its own set of apparatus. Lesson content was designed to teach specific components of the process of science. On the surface this setting appeared to require a departure from conventional teacher and pupil roles. A major aim of the study was to explore the extent of this departure and the variation in roles to be found in this setting. A further aim was to attempt to use pupil classroom behaviors to gain insight into the thought processes of pupils as they confronted the problems posed in this type of setting.

Theoretically, the rationale for the study derives from Barker's behavior setting concept and from Bernstein's concept of strong and weak framing as a way of viewing the range of options available to the pupil in the setting. One consequence of the behavior setting concept is that an individual's performance is a function of his role expectations and of his perceived goals of the setting. Pupil behavior might therefore be expected to differ in the classroom setting of interest here and, further, might be expected to vary with the degree to which the teacher attempts to directly control the activities of the pupils in this setting.

Grannis, for example, following Bernstein, has proposed that a high control setting (strong framing) is incongruent with a requirement for high pupil-pupil interaction and that such incongruence should influence performance. At the same time, aptitude-treatment interaction and matching model studies suggest that degree of congruence may be a function of pupil characteristics as well as of degree of teacher control.

The practical origins of the study can be found in a number of problems of implementing a process based science curriculum. The issue of teacher control was of interest in this context because most major elementary science curriculums do require at least a surface departure from the conventional classroom setting. There are indications that some of the difficulties in implementing such curriculums are related to this requirement. A further problem of implementation is that of measuring outcomes. An attempt was made in this study to use direct observation of pupil behaviors, combined with pupil interviews, to overcome some of the limitations of tests of science process.

The first phase of the study was developmental and descriptive in nature. Videotape recordings from thirteen science classes were used to develop a modified version of the Bellack classroom observation system. This system then provided the basis for a descriptive analysis of teacher-pupil and pupil-pupil interactions in the classroom. In particular, the behaviors of pupils in a group without direct teacher involvement were compared to those in the same group setting in the presence of the teacher. Teacher behaviors in the whole class and the group setting were also compared.

It was found that teacher behaviors showed remarkable stability in the class and group settings, and that these behaviors tended to conform closely to those found for move conventional class settings. Pupil behaviors, on the other hand, differed markedly depending on presence or absence of the teacher in the group setting. It was also found that much of the discourse was procedural rather than substantive in character. This finding limited the ability to use substantive discourse as a basis for analyzing the nature of pupils' thinking about science processes. It did, however, provide important insight into pupil roles in the classroom.

The second phase of the study involved an experiment on teacher control. Two experimental treatments were defined in terms of a number of teacher behaviors derived from the classroom observation. Ten grade six classes were exposed to a nine week sequence in each treatment. Teaching was carried out by four teachers assinged to the project, and trained to implement the treatments. Teachers were assigned to classes in a balanced fashion to reduce unique teacher-class or teacher-treatment effects. Treatments were monitored by videotaping each class on three occasions during the term, yielding about eight taped sessions for each teacher. These videotapes were also used as the basis for coding pupil behaviors.

Treatments, teachers, and a number of pupil characteristic variables were included as independent variables in the analysis. Criterion variables were achievement, preference, pupil interview items, and pupil classroom behaviors.

Results for the achievement criterion indicated no significant treatment effects. An interaction of treatment with pupil attitude did, however, appear. Strong treatment effects emerged for the two preference criteria. A treatment-teacher interaction and a number of interactions of treatment and pupil characteristics were also found. Pupil interviews tended to confirm the first phase finding of stable pupil and teacher roles. It was also found that pupils tended not to perceive the lessons in science process terms, although significant treatment and teacher effects did exist with respect to these perceptions. The analysis of pupil classroom behaviors showed treatment differences in the expected direction (e.g. more pupil talk in low control). Again, however, this result was complicated by the presence of treatment-teacher interactions as well as teacher main effects.

In behavior setting terms, these findings suggest that changes in teacher and pupil roles do not necessarily follow even from fairly drastic changes in the surface features of a setting. There is also some indication that role shifts may occur over time and that these role shifts ultimately can influence outcomes.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the active participation, support, and cooperation of the many people who made this study possible.

The collection of data for the first phase was made possible through the cooperation of the Bonavista-Trinity-Placentia Integrated School District, Clarenville, Newfoundland, the Avalon Consolidated School Board, St. John's, Newfoundland, and the Avalon North Integrated School Board, Bay Roberts, Newfoundland. Superintendents and principals were most helpful in granting permission and arranging visits to schools. The true burden, however, fell on teachers who willingly faced what to them must have appeared to be the prying eyes of cameras and observers. It is with pleasure therefore that we acknowledge the participation of:

1. Mr. K. Clarke
2. Mrs. A. O'Keefe
3. Mrs. J. Cooper
4. Ms. S. Blackmore
5. Ms. A. Robbins
6. Ms. G. Rickets
7. Ms. S. Carpenter
8. Mrs. D. Gullage
9. Mrs. L. Dalton
10. Mrs. P. March
11. Mr. G. Butt
12. Mr. B. Snook
13. Mr. C. Combden

The experimental phase required even greater intervention in the participating schools, placing extraordinary demands on principals, in particular. David Andrews, Neville Butler, Harry Eady, Thomas LaFosse, Frank Lee, and Robert Parrott were instrumental in making the arrangements to have their sixth grade science classes taught by teachers from the research team for a complete term. The teachers themselves, Rosetta Bishop, Judy Fagan, Joan Hiller, and Cynthia MacArthur also deserve special mention for their ability to translate definitions on paper into classroom operations.

Throughout the study, the Center for Audiovisual Education at Memorial University provided videotape equipment and technical services. Assistance in the supply of classroom materials was provided by the Science Education Section of the Department of Curriculum and Instruction at Memorial University.

The project received major financial support in the form of research grants from the Canada Council. The authors are most grateful for the continuing support of Canada Council for studies of the teaching process.

Most of the work on this report was completed while the first author was on sabbatical leave at Stanford University and Michigan State University. Special thanks are due to Professor N.L. Gage of the Center for Educational Research at Stanford and Professor Lee S. Shulman, Co-Director of the Institute for Research on Teaching at Michigan State for their efforts to ensure that the author's visit to these institutions was pleasant, stimulating, and productive.

Our final word of thanks goes to Judy Blundon for her excellent job of typing and organizing this report.

I THE PROBLEM

Objectives of the Study

This study was designed to address a number of questions of classroom dynamics and outcomes in the specific context of elementary science classes. The classes under investigation were characterized, on the surface, by direct pupil investigation of science problems, including pupil apparatus manipulation, data collection, and data interpretation. From the point of view of teacher-pupil relationships, the teacher typically interacted with the class as a whole at the beginning of a lesson, when a problem was introduced, and near the end, when pupil findings were summarized and an attempt made to achieve closure on the problem at hand. For the remainder of the lesson, the salient group for the study of interactions was usually a pair of pupils working together with the same apparatus. During this time, the teacher typically was circulating about the room, engaging in discussion with one group of pupils at a time. The classroom setting was thus, in a broad sense, similar to that which has come to be associated with many of the major elementary science curriculums.

More specifically, the objectives of the study may be stated as follows:

1. To develop a classroom observation system capable of yielding quantified data on both the substantive and instructional aspects of teacher-pupil and pupil-pupil relationships.

2. To determine whether such a system could be used to provide a measure of the nature of scientific thought processes of pupils during the ongoing activities.

3. To describe the range of teacher and pupil behaviors occurring in the classes and to compare behaviors in the teacher-class, pupil-pupil, and teacher-group settings, with particular reference to the question of degree of teacher control in the class and group settings.

4. To use categories of the observation system to define experimental levels of teacher control and to investigate classroom behaviors and outcomes under these experimental conditions, with particular reference to interactions between experimental treatments and pupil characteristics.

Rationale

The original impetus for the study was provided by a number of problems encountered in an earlier curriculum development project. As that project progressed to the implementation and evaluation phase, it became apparent that, in spite of the large amount of effort expended in producing "process" or "inquiry"[1] oriented science curriculums, very few attempts had been made to study the dynamics of the classroom setting required for the operation of such curriculums. Furthermore, the difficulties encountered in devising outcome measures appropriate to such curriculums led to the notion that it might be possible to shed some light on outcomes by directly observing pupils as they engaged in classroom activities. Finally, the issue of teacher control of classroom activities was considered worthy of investigation because, given the constraints imposed by a particular curriculum, it is only within the area of direction of specific activities (and perhaps, to some degree, in the choice of activities within the curriculum) that the teacher may be considered to have some freedom of choice. Although, on the surface, the new curriculums seem to require a relatively unstructured classroom environment, the unstructuredness in fact extends only to the requirement that the pupils carry out laboratory type activities. The degree to which the conduct of such activities (in terms of pacing, direction giving, narrowness of task, etc.) should be controlled by the teacher in a matter of teacher choice. Since variation in degree of teacher control is related to the question of increased or decreased pupil freedom to engage in a variety of activities in the classroom, and since pupil performance under different conditions of freedom might be expected to be related to pupil characteristics, it seemed desirable to investigate the effects of different degrees of teacher control on classroom behavior and on outcomes.

More theoretically, the rationale for the focus on teacher control derives from Barker's behavior setting construct (Barker, 1963) and from Bernstein's concept of strong and weak framing (Bernstein, 1971). More specifically, Barker has developed the notion that an individual's perception of the goals of a setting and of his role in the setting may influence his behavior. Thus, high and low teacher control may lead to different pupil expectations of goals and to different pupil role perceptions. For example, it might be hypothesized that in a high control setting the pupil may be more likely to perceive the goal in terms of meeting immediate teacher requirements, thus reducing the emphasis by the pupil on the cognitive demands of the problem at hand. Grannis (1973), following Bernstein, has proposed that classroom behaviors may be determined by whether a particular setting is congruent or incongruent in terms of teacher versus pupil control of a number of parameters of the setting. It might be argued that teacher control of lesson pacing, apparatus arrangements, results to be sought, and so on is incongruent with the requirement for pupil use of apparatus, pupil data collection, and high pupil-pupil interaction. In this study, the low control treatment may be interpreted as an attempt to bring about greater congruence in the various elements of lesson control.

The behavior setting construct admits the possibility that an individual's behavior can be a function of the individual's characteristics as well as of parameters of the setting. A more formal expression of this concept is given by Hunt (1971; Hunt and Sullivan, 1974) in the context of the question of matching teaching styles to pupil characteristics. Hunt's analysis is based on an elaboration of Lewin's paradigm, which views behavior as being a function of the person and the environment (B=f(P,E,)). It is this general notion, of course, which provides the rationale for the increased emphasis on aptitude-treatment interactions in recent studies of teaching.

The concept of control also relates to the "direct instruction" model proposed by Rosenshine (1976). Rosenshine's model is based on his analysis of the findings of a number of recent large scale correlational studies of the relationships between classroom behaviors and outcomes. These studies seem to suggest that direct teacher involvement in the instructional process (e.g. direct supervision of activities, high feedback, organization of time, etc.) leads to increased achievement in elementary reading and mathematics. Since this model seems to be in direct opposition to the prevailing wisdom in science teaching, with its emphasis on increased pupil independence, one wonders whether science teaching, particularly when directed towards process objectives, does, in fact, differ in some essential way from teaching in reading and mathematics.

From a slightly different perspective the study may be interpreted as an attempt to approach the investigation of teaching processes at the level of teaching "strategy". That is, the study attempted to define variations in teacher behavior at a level at which the teacher might ordinarily be expected to make decisions about these variations. This level may be distinguished from the "tactical" level at which behaviors are in terms of categories such as "praise" or "questioning" and from the "methods" level in which variations are more in terms of materials, curriculum, technologies and the like (e.g. studies of programmed learning versus conventional methods). In the latter case, the variations under investigation are likely to be outside the control of the teacher (i.e. a decision to adopt a programmed approach to teaching reading, for example, would likely be made at the school or district rather than the teacher level). Although our experimental treatments have been defined in terms of teacher behaviors, and thus it might appear that the study closely resembles a "tactics" study, it is the general combination of behaviors which identifies the notion of strategy.

Scientific Processes

As has already been pointed out, most of the elementary science curriculums produced during the era of the large scale curriculum development projects have emphasized, either explicitly or implicitly, the process aspect of science. In the broadest sense, this implies that curriculums have attempted to present science as a way of knowing and have attempted to deal with the various empirical and analytical processes that characterize the way in which scientific knowledge is acquired. Thus, in investigating science classes in which this type of curriculum is being used, it is necessary to have some model of what constitutes scientific processes and to deal with the problem of measuring process outcomes.

It is now recognized (c.f. Schwab, 1964) that earlier attempts to identify components of "the scientific method" resulted in an oversimplification of the dynamics of science. Nevertheless, in any philosophical analysis of the nature of science, the application of certain empirical and analytical processes in scientific investigation is acknowledged. Thus, for example, while the exact place of inductive and deductive logic in science may be debated, an understanding of the nature of these processes is essential in the practice of science. Similarly, it is generally recognized that before one can assert that a change in one variable is caused by a change in a second, it is necessary, as far as possible, to eliminate possible extraneous variables. The role of hypotheses as guides to investigation is acknowledged, even though the issue of whether scientists readily abandon hypotheses in the face of negative evidence is debatable (c.f. Kuhn, 1970). Classification is accepted as an initial way of imposing order on seemingly diverse phenomena and the logic of classification has been highly developed in some branches of science.

For purposes of this study, scientific processes are taken to be those operations identified by the American Association for the Advancement of Science in developing the curriculum Science - A Process Approach. The following slightly modified list of these processes was used in the Elementary Science Curriculum Study (Crocker, 1973):

Observing
Classifying
Quantifying
Communicating
Predicting
Inferring
Hypothesizing
Forming operational definitions
Interpreting data
Controlling variables
Experimenting
Formulating models

There is an obvious parallel between science processes in the AAAS model and the properties of logical thought as described by Piaget. In fact, it may be argued that the so called science processes are simply specific manifestations of a generalized form of rational thought. Piaget and his collaborators have studied in detail such processes as classifying, quantifying, and controlling variables in a developmental context (Inhelder and Piaget, 1958). In fact, Piaget equates the development of intelligence itself with the acquisition of ability to interact with the environment in successively more complex ways (Piaget, 1973, p. 10). Piaget further argues that such ability is acquired through successive encounters of the individual with phenomena in the environment. While Piaget's theory has often been used to identify limits in children's ability to use certain logical thought processes, it is also apparent that the full potential of a child at a particular developmental level is reached only if that child has had appropriate encounters with the type of problems that require the application of logical and empirical processes. In fact, it seems plausible to argue that the lack of such exposure may account for the failure of large numbers of individuals to reach the formal operations stage of development, even at ages well beyond those suggested by Piaget (Chiappetta, 1976).

The problem or devising measures of process outcomes in science teaching is one of the most difficult ones in the design of research in this field. One of the original reasons for initiating this study was to attempt to approach the measurement problem through classroom observation. Thus, a good deal of effort in the first year of the study was devoted to an attempt to determine whether evidence of science process thinking on the part of pupils could be obtained by coding categories and levels of science processes as the pupils engaged in interaction with each other and with the teacher in the classroom. As the study progressed, a number of modifications were made in this approach. Briefly, the attempt to define levels of thinking was considered unsuccessful. In the experimental study, therefore, alternative approaches were used, including pupil interviews and paper and pencil tests. The classroom observation data could still be used to examine the proportion of discourse devoted to science process categories but could not be used to identify levels of thought.

Overview of Method

The study was carried out in two phases, each of about one year in duration. In the first phase, videotape recordings were made of thirteen class sessions in which regular class teachers were engaged in teaching a science curriculum involving emphasis on scientific processes and requiring classroom activities by small groups (usually pairs) of pupils each with its own set of apparatus.[2] These videotapes were used in constructing and testing the classroom observation system and formed the basis for the descriptive analysis to be given in this report. The second phase consisted of an experimental study of teacher control. Ten grade six classes were taught in each of two modes (high and low control) defined in terms of frequency of occurrence of specific categories of teacher behavior in the classroom observation system. Four specially trained teachers were used to permit balanced assignment of teachers to classes and treatments. Order of presentation of the treatments was also balanced. A number of measures of pupil background, ability, personality, and attitudes were administered to allow the investigation of interactions. Dependent variables were pupil achievement in science processes, pupil preference for treatment, and a number of variables derived from a pupil interview schedule and from the classroom observation system.

The classroom observation system was derived from the system developed by Bellack, _et_. _al_. (1966). The basic structure of this system, including the "pedagogical move" concept and the division of discourse into substantive and instructional components was accepted. Modifications were made in the coding unit, the nature of the substantive dimension, and the categories of the instructional dimension. In addition, categories for the coding of physical actions were added. The coding structure was also changed to record both speaker and receiver of a discourse unit.

II RELATED RESEARCH

Teacher Control

Although a large body of research exists on the general issue of control, the problem has been viewed from many different perspectives, thus making comparisons difficult. For example, the large number of studies generated by the Flanders Interaction Analysis Category System[3] (FIAC) perceive the question in terms of teacher "indirectness". Kounin (1970) has investigated control in the sense of discipline and group management. Behavior modification studies (O'Leary and O'Leary, 1972) are concerned with control of deviant behavior. Control of classroom organization, grouping, materials handling, movement and the like have been treated in some of the large scale classroom observation studies (e.g. Soar, 1973; Brophy and Evertson, 1974; Stallings and Kaskowitz, 1974).

The conclusions that are reached about the cumulative findings of studies using the Flanders instrument seem to depend on the perspective from which the studies are reviewed. Flanders (1970), in his own review, reaches the tentative conclusion that increased "indirectness" is associated with increased pupil achievement and attitude scores. Dunkin and Biddle (1974), on the other hand, draw attention to the contradictions between individual studies. More recently, Gage (1976) has argued that

Dunkin and Biddle tended to rely too strongly on counting studies which did and did not find statistically significant differences rather than looking at the more general tendency displayed by the direction of the result across studies. By using a test of significance applied to the combined results of some twenty studies, Gage has shown that, taken together, the studies indicate that the relationship between teacher indirectness and pupil achievement is highly significant. The same procedure also yielded significant combined results for teacher praise and pupil attitudes, teacher acceptance of pupil ideas and pupil achievement, and teacher criticism and pupil achievement.

Kounin's research indicates that a number of "group management" variables are useful as predictors of the dependent variables of work involvement and deviant behavior. The variables of "withitness" (the teacher makes few errors in attempts to control behavior), "momentum" (the absence of behaviors that slow the pace of the lesson) showed particularly strong effects, with lesser effects being shown by "smoothness" (the absence of interferring behaviors) and "group alerting" (the involvement of many pupils in the task). A major limitation of the Kounin work, however, is the narrowness of the dependent variables, especially the lack of inclusion of outcome measures.

Somewhat mixed results have emerged from the more broadly based classroom observation studies. Soar and Soar (1973) found nonlinear relationships between variables associated with control and pupil achievement. Stallings and Kaskowitz (1974) suggest that greater "openness" is associated with higher scores on affective type of measures but that greater control is associated with higher achievement scores. Brophy and Evertson (1974) report interactions between control variables and socioeconomic status, with lower SES pupils benefiting generally from greater control. All of these results, however, are made difficult to interpret by the problem of deciding which variables, indeed, are reasonable reflections of the broad concept of control.

The specific aspect of control of interest in this study is somewhat narrower than most of those discussed above. To the casual observer, all classes in this study would perhaps appear to be characterized by low teacher control. That is, the surface features of these classes show pupils working in small groups with no direct teacher intervention for significant periods of time. Nevertheless, beyond this surface structure, it is possible to detect (and induce, in the experimental study) variations in the manner in which the problem is defined, the types of statements made by the teacher to pupils, the amount of direction giving, and so on. More specifically, in the experimental study, teachers were directed to vary a number of behaviors which could be identified by categories in the classroom observation instrument (these will be described in detail in a later section).

Thus, it was possible to develop at least a quasi-quantitative approach to variation in teacher control.

A number of recent studies have a more direct bearing on the problem of interest than do the studies cited above. Shymansky and Matthews (1974) report an experimental study on two learning conditions in fifth grade science. The two learning conditions, referred to as "teacher structured" and "student structured" were defined in terms of a number of observable teacher behaviors related to teacher directiveness. Student classroom behaviors and student performance on a science process test were used as criterion variables. Results showed that student behaviors differed markedly under the two experimental treatments. Surprisingly, almost no behavior differences occurred as a function of differences in student ability. Interactions between student ability and treatment were also generally nonsignificant. With respect to science process performance, the student structured treatment led to significantly higher scores. In terms of the present study, it is also interesting to note that the amount of time spent by the teacher in interacting with individual students correlated negatively with the amount of time the student spent on activities of his own design. This suggests the need for a closer examination of teacher-student interactions.

A similar, but smaller scale, study by Crocker, Bartlett, and Elliott (1976) yielded results in the opposite direction from those of Shymansky and Matthews. Students in a more highly structured learning environment performed better on a science process test. In this study, all subjects were exposed to both treatments so that it was possible to obtain a measure of preference for treatment. Again, the results seemed to favor the structured treatment.

Differences in the results of these two studies can possibly be accounted for by differences in methodology. Since the Shymansky and Matthews study was of much longer duration, it is possible that improved performance in an unstructured environment occurs only after a long period of orientation to this type of environment. The strongest effects in the Shymansky and Matthews study were, in any case, not in student performance but in classroom behaviors.

A study that even more closely parallels the experimental phase of the present study has been conducted at the Stanford Center for Research and Development in Teaching (Program on Teaching Effectiveness, 1976). In this experiment, a three way factorial design, based on two levels each of teacher structuring, soliciting, and reacting (Bellack, et. al., 1966) was employed. Teacher behaviors on each of the factors were controlled by using scripted lessons. In general, the results indicate that low soliciting was more effective in bringing about achievement on test items for which the teacher was the only source of information.

High structuring and high reacting were also more effective than their low counterparts. Thus, the treatment combination of low soliciting - high structuring - high reacting was the most effective for this criterion. No differences between treatments occurred, however, on an essay achievement test and an attitude scale. A generalized regression analysis indicated, as might be expected, that pupil ability accounted for more of the variance in achievement than did treatment effects. It was also found that ATI effects accounted for about the same proportion of variance as treatment effects. Nc conclusions about specific sources of ATI's were reported.

The present study was intended to replicate the earlier Crocker, Bartlett, and Elliott study and to incorporate a number of design improvements over that study. In terms of the Program on Teaching Effectiveness (PTE) study, to the extent that our high control treatment and the PTE high structuring - low soliciting - high reacting treatment reflect the underlying notion behind Rosenshine's "direct instruction" model, there is reason to predict that, at least for achievement, the high control treatment should be more effective (a result which would also be consistent with Crocker, Bartlett, and Elliott). At the same time, our process achievement measure is somewhat different from the PTE measure of knowledge so that differences may occur as a result of differences in criterion measures. In the present study, the use of a repeated treatments design permits a measure of preference to be developed and also permits what amounts to two replications of the same experiment in a single study, thus reducing the likelihood of obtaining false positives as a result of the large number of statistical tests used.

Small Group Interaction in the Classroom

The area of pupil-pupil relationships seems to have been as widely ignored as teacher-pupil relationships have been widely studied. Perhaps the reason for this can be found in the observation that pupils in the typical classroom must behave as if they were in solitude, even when, in fact, they are surrounded by other individuals (Jackson, 1968). The assumption, perhaps well founded, in most classroom studies is that the pupil interacts only with the teacher, even though, under any but classroom conditions, he might be expected to engage in many forms of interaction with his fellow pupils.

In the type of science classes under investigation here, pupil-pupil interactions constitute an important feature of lesson structure. The requirement that pupils work in pairs, each with its own set of apparatus may be regarded as one of the primary boundary conditions influencing classroom behaviors in this setting. This grouping structure dictated that most pupil-pupil interactions would be dyadic in nature and would occur within specific pairs. The pairwise interaction pattern was, of course, commonly broken by teacher interactions with the entire class and by teacher interventions in specific groups. Less often, inter-group interactions tended to occur or else group structures became more diffuse.

In spite of the pervasiveness of small group work in the science laboratory setting, the influence of group structure has been investigated in only a small number of studies. Hurd and Rowe (1966) identified compatible and incompatible groups based on a theory of compatibility developed by Schutz (1958). They then used the Bales (1950) observation system to investigate group behavior. Mayfield (1976) used Bales' (1970) model of group structure to identify group types. These types were compared with respect to the nature of the group solution to a science problem. In a more highly controlled study, Abraham (1976) established six groups based on three levels of divergent thinking and two group structures (homogeneous and heterogeneous) and explored verbal interaction patterns in these groups using categories of the Aschner-Gallagher observation system (the categories of this system are based on levels of cognitive functioning rather than on social-emotional climate).

While these studies are not directly comparable, they do suggest that group structure is significantly associated with in-group behavior and with outcomes. Unfortunately the studies do not clarify the problem of which group structure variables offer the greatest discriminatory power. In this study, a more exploratory approach is taken in that data on a large number of pupil characteristics and classroom behaviors were available. Also, the use of groups of two rather than the larger groups used in the studies cited should simplify problems of interpretation of the influence of one individual on another. The design of the present study also permitted the investigation of instances in which the teacher intervened in the work of small groups. In fact, the findings for teacher intervention are regarded as among the most important outcomes of the study.

Pupil Characteristics, Classroom Behaviors, and Outcomes

A reasonably large body of research exists on the relationships between pupil characteristic variables and the nature of teacher-pupil relationships. Variables such as pupil sex, socioeconomic status, race, ability, personality, physical appearance, and speech characteristics have been found to relate to teacher differential treatment of pupils in the classroom (Brophy and Good, 1974). More specifically, studies of teacher expectation effects seem to leave little doubt that teacher expectations do influence behavior, although the specific circumstances under which expectation effects occur remains the subject of some controversy (Brophy and Good, 1974; Snow, 1974).

A second area of controversy related to pupil characteristics centers on what has come to be known as aptitude-treatment interactions (ATI's). Bracht and Glass (1970), after reviewing the evidence to that time, reached highly pessimestic conclusions about the existence of such effects and the value of research in this area. A somewhat more positive view, however, has been expressed by Berliner and Cahen (1973). Hunt (1975) has examined, and rejected, a number of reasons for resistence to the adoption of an interactive model and illustrated how the broad B-P-E model can be developed in educational context with the environmental variable "degree of structure" and the person variable "conceptual level". While Hunt's work provides an excellent example of a theoretical approach to deriving potentially fruitful interactive effects, each such example yields possible relationships among only, at best, a small number of variables, and few such examples exist.

More recent reviews of empirical studies of ATI's have been made by Solomon and Kendall (1976) and Tobias (1976). Solomon and Kendall concentrated on the issue of "open" versus "traditional" classrooms and concluded that reason seemed to exist to puruse a major study in this area. Tobias was concerned with the apparent content specificity of ATI's and the possible shift in "degree of instructional support" required by a student at different stages. Tobias thus proposed that "prior achievement" in the subject at hand be considered the important variable in terms of interaction with instructional support (Tobias' term instructional support seems similar to Hunt's "degree of structure" and has some parallels with our concept of control). A number of studies were cited which tended to support the prior achievement hypothesis.

Solomon and Kendall's (1976) study is perhaps the most elaborate example of an empirical approach to the study of interactions. Solomon and Kendall measured a large number of pupil characteristics and used factor analysis followed by cluster analysis on the factors to reduce these to three major pupil clusters. A similar procedure for classrooms, based on classroom observation data, yielded six classroom clusters. Solomon then proceeded to examine interactions between pupil types and classroom types with respect to a number of dependent variables. Solomon's results are highly complex but are perhaps best summarized by considering the classroom environmental influences on the three types of pupils. Relatively unmotivated, low achieving pupils tended to perform better in permissive, varied classrooms and those which combined warmth with an emphasis on expressiveness. High achieving, highly motivated children, on the other hand, performed better in classes characterized by high control and moderate to high opportunity for student initiation. Noncompliant, highly self-directed children showed mixed results on the dependent variables, with activity and curiosity being maximum in permissive classrooms while creativity was maximized in more controlled classrooms. In most cases, further interactions with sex were found to exist.

The Brophy and Evertson (1974) study, while not specifically a study of interactions, did investigate the relationship between classroom environments and outcomes in low and high SES schools. Variations in the correlations found for the two types of schools strongly suggest interactions between classroom environment and SES. Theoretical support for the existence of this specific type of interaction is found in Bernstein's (1971) concepts of "linguistic codes" and the relationship of such codes to social status. Bernstein's work suggests that SES may simply be a proxy variable for a more underlying trait based on language experience. This underlying trait may also account for correlations between SES and academic ability and may thus also explain any interactions that might exist between ability variables and classroom environments.

Although the results of recent studies seem to support an interactive model, the problem of selecting potentially useful variables remains, as does the issue of whether interactions are content specific and outcome specific. In this study, the use of a variety of dependent variables suggests that a range of pupil characteristic variables should be investigated for interactions. Intuitively, one would expect that academic ability characteristics would be more likely to yield interactions with respect to achievement outcomes, while personality related variables might yield interactions for classroom behavior and preference type outcomes. The selection of pupil characteristics used was based on this notion and on the results of specific studies in which classroom environments were varied in a manner reasonably consistent with the variation in control used in this study.

III METHOD

Samples and School Settings

As already indicated, the study took place in two phases, the first exploratory and descriptive and the second experimental. Two distinct samples were therefore used.

The first phase sample consisted of thirteen classes selected from one urban and two rural school districts in Eastern Newfoundland. Six of the classes were second grade, three were fourth grade, and four were sixth grade. All sixth grade teachers were male, while all those in the lower grades were female. Since the first phase was essentially exploratory, no specific selection criteria were applied other than the requirement that the classes be using the particular science curriculum of interest, and that the teacher be willing to have videotaping carried out in the class. However, some standardization of lesson content was imposed in order to reduce the influence of differences in lesson types.

The original experimental sample consisted of eleven sixth grade classes in six schools within the Avalon Consolidated School District in St. John's, Newfoundland. One of these classes was subsequently dropped from the main analysis when it was found that pupils had already completed some of the course content covered in the experimental program. This class was then used for a number of other purposes in the study, particularly in reliability studies on some of the instruments.

The region from which the samples were selected has a population of about 250,000, of which about half live in the St. John's metropolitan area and the remaining half in a large number of small communities, nearly all with populations under 5,000 and many with populations under 1,000. Although the region is not highly industrialized, a wide range of occupations exist, even in the rural areas. One of the regions consists of a few heavy industries combined with the traditional fishing based occupations. The urban area is largely a government and university center with little primary industry. A wide range of social and occupational classes thus exists throughout the region, although the predominant group might be appropriately labelled "rural working class".

The school systems in this region, particularly those in the rural areas, have been undergoing a transition from small community based schools to larger units serving several communities. Many children, even in lower grades, travel by bus. Schools in the first sample were typically K-6 or K-8 schools with one class per grade level. In the second phase sample, the schools were generally larger, ranging in type from older, central city schools to new suburban schools with open areas, resource centers and similar features. Perhaps surprisingly, it was the suburban schools that tend to be most crowded while the central city schools appeared to have surplus space because of declining enrollments.

The curriculum used in the schools was relatively standardized because curriculum decisions are generally made at the provincial level. All schools in the first phase sample were using the ESCS science program (Crocker, 1973). Schools in the experimental sample appear to have been teaching little science prior to the introduction of the experimental program, with the notable exception of the class that was subsequently eliminated from the analysis.

The Curriculum

As already noted the science program used was a locally developed one structurally similar to the AAAS program Science - A Process Approach (AAAS, 1968). Emphasis at the second grade level was on the processes of observing and classfying. Quantifying was perhaps the process most strongly emphasized at the grade four level, while interpreting data and controlling variables were characteristic of the grade six level. For the experimental study, the curriculum was modified somewhat both to facilitate the experimental treatments and to accommodate the fact that pupils in the experimental sample had not had previous exposure to this type of program. In particular, it was necessary to pursue the process of quantifying in the experimental version of the grade six program because it could not be assumed that subjects had acquired prerequisite skills in this area.

Most activities in the program were designed to occupy one or two class sessions, with apparatus provided to permit pupils to work in groups of two. An activity would typically commence with a teacher introduction, continue with a data collection stage in which pupils would have to set up apparatus and make measurements of some sort, and end with a class discussion based on the data gathered. This general pattern was followed in the experimental treatments. The manner in which the treatments were differentiated is described in a later section.

Data Collection - Descriptive Study

Videotape recordings of a single class session in each of the thirteen classes formed the basic data for the first phase of the study. While this may appear to be a relatively small data base, particularly in that teacher behavior across occasions was not sampled, it must be emphasized that the purpose of this phase of the study was not to compare teachers but rather to describe pupil behaviors (as well as to develop and test the category system). In these terms, the size and representativeness of the pupil sample is of more interest than the class sample. Teacher differences were reduced to some degree by standardizing the lesson content and the general format of the lesson (the latter was dictated to a large degree by the nature of the curriculum).

In each classroom, two independently operated portable VTR units were used, each of which was directed towards a randomly selected pair of pupils working as a group. This process yielded usable data on fifty pupils. Teacher verbal behaviors were recorded when the teacher was interacting with the whole class, while both verbal and nonverbal teacher behaviors could be recorded when the teacher interacted with one of the groups of pupils selected. This allowed descriptions of teacher behavior to be developed and the question of intra-class compared to inter-class variations in pupil behavior to be explored.

In addition to the video recordings, data on pupil academic ability, reading ability, personality, sex, and self concept were gathered to permit preliminary exploration of the relationships between these characteristics and pupil behavior, particularly in the context of whether pairs of pupils were matched or unmatched on these characteristics. Detailed descriptions on the specific measures used are given in the chapter on instruments.

The actual coding of classroom behaviors was carried out by first preparing typed transcripts of the audio component of the recordings. Each sentence was then coded using the categories to be described in the next chapter. Non verbal behaviors were finally coded directly from the videotapes, using the transcripts as a guide.

Experimental Design

The general concept of control was expressed in operational terms using a number of variables derived from the classroom observation system used in the first phase. The basis for choosing these variables was that they appeared to provide relatively low inference measures of the amount of time the teacher spent directing the activities of the pupils. High and low control treatments were defined in terms of these variables in a quasi-quantitative manner. That is, teachers were directed to engage in behaviors representing high or low values of the variables concerned but were not specifically instructed as to a specific range of values of these variables appropriate to each treatment. The latter approach was judged not to be possible without detailed scripting of the lessons, a practice that would have tended to reduce the representativeness of the lesson format.

The variables used, and their planned relative values in the high and low control treatments are given in Table 1. Although the values of these variables were not dictated in advance, it was possible to monitor the teacher's adherence to the treatments by observing her behavior on these variables as lessons were taught. Some feedback was given teachers as the treatments progressed, although this was not as extensive as would have been desired because detailed coding of the variables was not completed in time.

Table 1

Summary of Variables Used in Differentiating Experimental Treatments

Variable	Relative Value	
	High Control	Low Control
Teacher-Class Interaction	High	Low
Teacher-Group Interaction	Low	High
Proportion of Pupil Talk During Teacher-Class Interaction	Low	High
Proportion of Pupil Talk During Teacher-Group Interaction	Low	High
Teacher Structuring	High	Low
Teacher Soliciting	Low	High
Teacher Responding	Low	High
Teacher Reacting	High	Low
Ratio of Commands to Total Teacher Solicits	High	Low
Ratio of Requests to Total Teacher Solicits	Low	High
Pupils Reporting Individual Results	Low	High
Teacher Substantive-Logical Discourse	Low	High
Pupil Substantive-Logical Discourse	Low	High
Average Wait Time (Sec)	Short	Long

From the beginning, it was expected that differences between teachers would occur and would reduce the uniformity of application of the treatments. At the same time, such differences would increase the generalizability of the findings. This assumed, of course, that the differences were not sufficiently large to obliterate differences between the treatments themselves. Data on differences between treatments and between teachers are presented in the chapter on data analysis.

Four experienced elementary teachers attached to the project staff were given a one week training program on the treatments and the general outline of the lesson content. One of these teachers was engaged full-time to act as coordinator for lesson preparation, distribution of materials, sequencing, and related tasks. The remaining teachers were part-time. All teachers were free to be assigned to classes in accordance with the experimental requirements. Regular classroom teachers were replaced by these teachers for all science periods during the eighteen-week course of the experiment. In this case, the loss in generalizability stemming from the use of special teachers was offset by the substantial gain in experimental control of being able to assign teachers to classes in a balanced fashion, avoiding the problem of unique teacher-class effects that are well known in studies using intact classes.

Because of the relatively small number of classes, it was decided to employ a repeated treatments design in which all classes were exposed to both treatments. At the beginning of the experimental period, five classes were assigned to low control and six to high control. At the end of nine weeks, the treatments were reversed and the teachers reassigned to new classes. The assignment of teachers to classes was random, with the restriction that each teacher have at least one class in each treatment. Scheduling restrictions imposed some constraints on teacher assignments. However, these constraints were not considered relevant to any of the experimental variables and were therefore not considered departures from randomness.

On the whole, the design permitted the analysis of treatment main effects and treatment by student characteristics interactions free from confounding due to teacher and class effects. Some confounding of treatment-class and treatment-teacher effects did exist within a round. However, this confounding was reduced by treating the two rounds as replications, with teachers being assigned to different classes in the two rounds. It should be noted, of course, that the second round could not be a complete replication of the first because of the possibility of carry-over effects from the first treatment to the second. However, the duration of each treatment, and the fact that the criterion tests were specific to the treatment or to the content just completed tended to reduce this problem.

Data Sources - Experimental Study

This section describes briefly the specific variables on which data were gathered and identifies the instruments used. An outline of the data collection sequence is also given. Details on the instruments themselves appear in the chapter on instrumentation.

The following is a summary of variables and the names of the various instruments.

Independent Variables

Experimental Treatments

As described in the preceeding section - implementation of the treatments was monitored using the classroom observation system.

Pupil Background Variables

Age

Sex

Socioeconomic status (Blishen Scale)

Years of science experience } Pupil questionnaire

Type of science experience }

School

Class

Pupil Abilities

Non-verbal intelligence (Raven Progressive Matrices)

Basic Skills (Canadian Test of Basic Skills,
Tests L.W.M. and composite score)

Reading Ability (Gates-MacGintie Reading Test,
Survey D, Form 2)

Personality and Related Skills

Extraversion } Junior Eysenck Personality Inventory
Neuroticism }

Attitudes toward school and science

Dominance-Submissiveness (Children's Personality
Questionnaire)

Dependency (Dependence Proneness Scale)

Self Concept of Ability (modified Brookover Scale)

Responsibility for success (Intellectual Achievement
Responsibility)

Dependent Variables

Science Process Achievement

Tests of measuring, controlling variables, and interpreting data developed for this study used as pretest and first and second round posttests.

Pupil Preference and Perceptions of Treatments

Student Preference Instrument - 22 item scale based on variables used to define differences between treatments.

Choice Task - Pupils asked to choose a science activity in either high or low control mode.

Interview Schedule

Pupil Classroom Behavior

A number of variables derived from the classroom observation schedule, coded from videotapes of sample classes.

In order to avoid subjecting pupils to a heavy testing load at the beginning of the study, data on the various independent variables were gathered at different times throughout the study, with the restriction that a particular instrument be administered to all classes at about the same time. This restriction was necessary in order to reduce the possible effects of shifts on these variables due to passage of time. It had originally been decided not to administer tests of basic skills because of the amount of testing time required. However, late in the school term, the school district decided to administer the Canadian Test of Basic Skills to all grade six pupils. Data on basic skills were therefore obtained at no cost to the study and with minimum interference with the experimental treatments.

The Gates-MacGintie reading test had been administered in two of the schools prior to the commencement of the study. In order to make these scores comparable to those for the remaining schools, scores were projected to the time at which the remaining tests were administered.

The science process achievement test was administered as pretest and first and second round posttests respectively (specific differences in the versions of the tests for each round are described in the next chapter). The preference instrument was also administered at the end of each round, with pupil responses being scaled with respect to the treatment just completed. Similarly, at the end of each round, pupils were asked to choose the mode in which they wished to carry out a particular activity (this activity was designed specifically for the choice task and was unrelated to the treatment activities themselves). In order to make the choice task realistic, pupils were required to carry out the activity in the mode chosen. Thus the choice task occupied a complete class session, even though the data were complete as soon as the choice had been made.

Originally, it had been planned to conduct pupil interviews at the beginning and end of each treatment. However, time constraints at the point of transition between the treatment rounds made it necessary to restrict the final Round 1 and initial Round 2 interviews to half the classes. All interviews were conducted immediately following a science class session, with interview questions being asked in the context of that session. In each case, a sample of eight pupils (four pairs) was selected to be interviewed. The interview lasted about five minutes for each pupil. Usually, a team of two interviewers was used in order to reduce the amount of time that classes would be disrupted by pupils being interviewed.

The videotaping procedure was modified somewhat in the experimental study in order to obtain a more representative sample of teacher behavior and to broaden the sample of pupils within each class. Essentially, the modification consisted of selecting five pupil pairs, plus the teacher, as subjects for recording in a single class session. Each VTR unit recorded three groups (three pupil pairs or two pupil pairs plus the teacher) by taking successive three-minute samples from each group throughout the lesson. A single thirty-minute tape was used for each class session, which permitted three to four three-minute observations for each group. Gaps of appropriate length were left between the three-minute recordings if the lesson was longer than thirty minutes in order to ensure that behaviors throughout the lesson were sampled (that is, to ensure that the tape did not run out before the end of the lesson, resulting in a lack of data on behaviors during the lesson summary phase).

Data Analysis

Descriptive Phase. The basic data for the descriptive analysis consisted of frequency tabulations for the various behaviors derived from the classroom observation system. These frequencies were transformed to proportions to permit direct comparisons between pupils, between groups, and between classes. The proportions were also used to develop a profile of teacher and pupil behaviors. Correlations were computed between pupil characteristic variables and classroom behaviors. Contingency tables were prepared to permit comparisons of pupils with different characteristics and to investigate groups who were matched or mismatched with respect to specific characteristics.

Episodes in which the teacher interacted with one of the groups being observed (referred to as teacher interventions) were analyzed separately from other aspects of the lesson in order to shed light on the general question of whether conventional teacher-pupil roles are maintained in the small group setting. In particular, comparisons were made between teacher and pupil behaviors in the intervention setting, teacher behavior in the whole-class setting, and pupil behaviors in the group setting with the teacher absent.

Experimental Phase. One major problem in studies using intact classes is in the choice of the class or the individual as the unit of analysis. The argument for the class as the appropriate unit rests on the assumption that observations made on individuals within a class are nonindependent. The argument for the class as the unit of analysis, perhaps best summarized by Cronbach (1976), rests on the assumption that observations made on individuals within a class are nonindependent. At the same time, to the extent that this assumption is invalid, there is a substantial loss of degrees of freedom when the class is used. In the present study, since much of the work in classes is at the group rather than the class level, the problem of class effects is not considered to be as great as is usually presumed to be the case. Furthermore, since regular classroom teachers were not used and since the project teachers taught both treatments, the usual coherence of the class-teacher unit was disrupted and unique teacher effects could be balanced in the treatment comparisons.

These considerations led to the choice of the individual pupil as the unit of analysis. Nevertheless, since the classes did differ in initial ability, and since with a small sample of classes there could be no assurance that these differences would balance out in the random assignment to treatments, adjustments for initial ability were incorporated into the analysis for the achievement criterion. Such adjustments were not made for preference since preference scores were uncorrelated with any of the ability measures used.

The basic analysis technique used was a generalized regression analysis. Such an analysis differs from the conventional analysis of variance or covariance in that it permits the assessment of ATI effects without blocking on the aptitude variables. This solves the problem of loss of power associated with blocking (Cronbach and Snow, 1977). This form of analysis also permits the investigator to choose the order in which the various effects are to be entered, in accordance with his overall model. In this case, variables were entered in the following order: (a) initial ability (IQ and CTBS composite scores entered simultaneously), (b) teachers, (c) treatment-teacher interactions, (d) ATI's with each attribute variable entered in stepwise order, (e) treatments. The argument for entering treatments last rests on the basic assumption of ATI studies that knowledge of treatment effects is less interesting than knowledge of how treatments interact with individual characteristics. The stepwise entry of the attribute variables was designed to yield an assessment of the relative contribution of each variable. It should be noted that although a pretest was used for achievement this pretest was not entered in the analysis because it was found that it made no significant contribution to posttest variance after initial ability effects were accounted for.

Because the choice task yielded only nominal data, the analysis for this variable consisted of a series of chi-square tests based on frequencies of choice for each treatment tabulated for each student attribute variable separately (in this case there was no feasible alternative to blocking on the attribute variables). In general, this procedure yielded results that could be directly compared to those for the remaining criterion variables, with the exception that the various ATI effects could not be assessed cumulatively.

IV INSTRUMENTS

This chapter describes in detail the various instruments used in the study. For standardized tests and other instruments adopted from external sources, the descriptions are based on published information on these instruments. For those instruments developed specifically for this study, the process of development is described and issues of validity and reliability discussed where appropriate.

Pupil Background Variables

A Socioeconomic Index for Occupations in Canada (Blishen Scale). Procedures used in developing this scale are described by Blishen (1967). Originally, the scale was derived by computing combined standard scores for education and income for some 300 occupations, based on the 1951 Census of Canada. A later revision extended to the full range of occupations included in the 1961 Census a set of ratings of prestige of occupations through the development of a prediction equation based on the prestige ratings and the education-income scores.

For purpose of this study, information was obtained from school records on the father's occupation for each pupil. Occupations were then grouped into seven categories on the basis of the Blishen Scale. These seven categories were used as attribute scores in the various analyses.

Previous Science Experience. Because previous experience in the type of science program used in the study was judged relevant to performance in the study, and because schools differed widely in the amount and type of science taught previously, pupils were asked to respond to two items designed to yield measures of previous experience. The first item dealt with the grades in which they had been taught science, while the second required the selection of categoreis describing type of science experience (whether activity or textbook based). These items yielded two variables - type of science experience and years of science experience.

Pupil Ability

Raven Progressive Matrices. The Raven Progressive Matrices (Raven, 1960) yields a measure of non-verbal intelligence. The scale has a test-retest reliability of from .83 to .93 depending on age of subjects. It was decided to use this test rather than a more comprehensive test of intelligence because other measures were available which were considered to be sufficiently well correlated with verbal intelligence to justify dispensing with the usual verbal intelligence measurement. Non-verbal intelligence was also considered to be important to the study because of the nature of the science activities used.

Gates-MacGintie Reading Test. This test series is one of a number of widely used standardized reading tests. The test was selected largely because the original information available had suggested that scores would be available for all pupils. However, it was later found that only two schools had data available for the current year. It was therefore necessary to administer the tests in the remaining schools and to make appropriate adjustments for differences in times of administration (in order to have controlled the time of administration, it would have been necessary to give the test a second time to pupils in two schools. This was judged less desirable than the adjustment procedure).

Survey D, the particular version used, has reported reliabilities (split-half) of .89 and .95 for the vocabulary and comprehension subtests respectively.

Canadian Test of Basic Skills. This test battery is a modified version of the Iowa Test of Basic Skills. The tests measure general intellectual skills in vocabulary, reading, language, work study, and mathematics. Because separate reading tests had been used, only the language, work study, and mathematics scales and the composite of these scores were used. These tests had been administered by the school district late in the term during which the study took place.

Split-half reliabilities for the various subscales of CTBS are reported to be in the range of .87 to .97.

Personality and Perceptions

Junior Eysenck Personality Inventory. The two traits measured by this instrument, extraversion and neuroticism, were judged relevant to the study because each has a bearing on how a person behaves in his relationships with others. There is also evidence that anxiety or neuroticism exhibits interactions with teaching styles.

Construct validity of the instrument is based on the nature of the intercorrelations between various personality traits. Split-half and test-retest reliabilities for the scales are in the .7 to .8 range. The inventory contains a Lie scale designed to detect the faking of responses.

Children's Personality Questionnaire. This instrument is a derivative of Cattell's 16PF (Porter and Cattell, 1968) designed for use with young children. For purposes of this study, only the submissive-dominant subscale was used. This scale was selected because it was judged relevant to the issue of the observed imbalance in student behavior in some student groups, with one student tending to dominate the activity. The question of interest was whether dominance of activity was related to the broader personality characteristic measured by the submissive-dominant scale and whether this characteristic was ultimately related to performance.

Test-retest and parallel forms reliability coefficients for

Factor E of the CPQ (of which submissive-dominant is one of four subscales) are in the range .74 to .82.

Dependence Proneness Scale. This scale was developed by Flanders, Anderson, and Amidon (1961) and consists of 45 agree-disagree items. The scale was used because variations in teacher control was obviously related to degree of pupil freedom to make his own decisions during the course of the activity.

The original authors reported a reliability coefficient of .68. A subsequent reliability test on grade six students carried out by Elliott (1973) yielded a coefficient of .91 (10 day test-retest).

Self-Concept of Ability. The scale used for this variable was a modified version of the Brookover (1962) questionnaire. The scale consisted of six five-choice items on the aspect of general self-concept of ability. Test-retest reliability was established at .72.

Intellectual Achievement Responsibility (IAR). This scale, developed by Crandall, Katkovsky and Crandall (1965) is aimed at assessing pupils' beliefs concerning internal versus external responsibility for intellectual-academic achievement. In addition to the total I score (internal responsibility), the scale is constructed to yield separate scores for internal responsibility for success (I+) and failure (I-).

Split-half and test-retest reliabilities for the instrument were in the .60 - .74 range.

This construct was used because of the expectation that individuals who tended to take responsibility for their own successes and failures would perform relatively better in the low control teaching mode.

Attitudes to School and Science (ATSCH and ATSCI). Pupil's attitudes to school and science were measured using a ten scale semantic differential on the objects "school" and "science" using adjective pairs from the evaluation dimension of the semantic space (Osgood, *et*. *al*., 1956). Test-retest reliabilities for the school and science subtests were .83 and .74 respectfully.

Science Process Achievement

Three twenty-item tests of science processes were developed for use as pretest and first and second round posttests. Items for these tests were adapted from the competency measures of Science - A Process Approach (AAAS, 1968), the AAAS Science Process Instrument (AAAS, 1970), and the Test of Science Processes by Tannenbaum (1971). For purposes of this study, the processes tested were confined to measuring, controlling variables, and interpreting data.

An initial pool of forty items for each test was used to prepare four parallel forms, with partial overlap of items between forms. The four forms were designed to overcome bias due to the shortness of each test and the possible uniqueness of items. The forms were randomly administered to students within each class. Items were designed both to reflect the immediate content of the curriculum and more general applications of the processes of interest. Item analysis procedures were used to obtain data on the effectiveness of each item. The tests were also pilot tested on subjects not otherwise involved in the study to ensure that the reading level was appropriate.

Reliability data were obtained using pupils in a school in which the same material as in the study was being taught but which was not used in the data analysis. Reliability for the pretest was found to be .69 while the posttests yielded reliabilities of .74 and .82 respectively.

Pupil Perceptions and Preferences

Pupil Preference Instrument. The starting point for constructing this instrument was the set of variables used to distinguish between the experimental treatments (see p.). For each variable one positively worded and one negatively worded statement was prepared in which pupils were asked to indicate whether they preferred the teacher to engage in more or less of the behavior indicated by each variable. For example, pupils were asked whether they preferred the teacher to give more or fewer directions or engage in more or less teacher-class interaction.

This instrument was administered at the end of each treatment round so that in each case pupil responses were made from the perspective of the treatment just completed. In deriving a score on the instrument, it was necessary to scale responses according to the treatment. It was also necessary to consider responses to a pair of items (negative and positive on the same teacher behavior). Four possible response patterns could occur for each item pair (prefer more control, prefer less control, prefer same amount of control, and inconsistent). By summing over all item pairs, and adjusting for inconsistent responses (e.g. it was inconsistent for a subject to respond that he preferred both more and less teacher direction), it was possible to derive a score representing the student's liking for the treatment he had just completed.

Test-retest reliability was established at .71.

Choice Task. As an independent, more task-oriented, measure of pupil preference for treatment, pupils were asked, at the end of each treatment, to choose a particular activity to be taught by either of two methods, corresponding to the differences between the treatments. Essentially, the choice given was between being assigned an activity in which the purpose was identified and apparatus supplied but no specific directions given as to data collection, procedure, etc. and the same activity in which a worksheet was supplied giving detailed directions for the activity.

Data from this task consisted simply of recording the choice made by each pupil. Once the choices were made, pupils were permitted to carry out the activity in the manner chosen, in order to ensure that the choice was based on a genuine expectation that the task would have to be carried out.

Interview Schedule. The interview schedule was originally designed to achieve a number of purposes, including the provision of an independent measure of process achievement using an assigned task. Because of time limitations, however, the latter objective could not be achieved. Nevertheless, the interview yielded data on whether subjects perceived the lessons in process, content or procedural terms, whether they preferred greater or less teacher control, and on the pupil's relationships with his partner.

Essentially, the interview was a relatively highly structured one in which pupils were asked specific questions about the purposes of the activities, the degree of teacher control, and pupil-pair relationships. Pupil free responses were categorized by the interviewer according to a predefined set of categories. Provision was made for write-in of pupil responses that could not be classified by the original scheme.

Reliability of this instrument was approached both from the point of view of stability of pupil responses and of interviewer classification of these responses. The first was accomplished by a re-interview of pupils following an activity which relatively closely approximated the conditions of the original interview. Interviewer stability was assessed using audio tapes of a set of interviews. Both inter-interviewer and intra-interviewer stability could be approached in this manner.

Classroom Observation System

Because the development of this instrument was one of the objectives of the first phase of the study, and because of its importance to the remaining aspects of the study, this instrument is described in somewhat more detail than those above.

The basic requirements of the system were as follows:

1. Behaviors related to both the substantive and instructional aspects of the classroom activities had to be recorded.

2. The specific substantive behaviors of interest were those that might be indicative of pupils' ability to use various science process skills.

3. Instructional behaviors of interest were those related to the concept of control. More specifically, teacher behaviors should be those that showed the distinction between those teachers who exercised a high degree of

control over the moment-to-moment activities of pupils and those who allowed greater pupil freedom.

4. The system had to be capable of identifying individual pupils.

5. Pupil-pupil as well as teacher-pupil behaviors were to be recorded.

6. Physical actions, as well as verbal utterances, were to be recorded.

A review of classroom observation systems available at the beginning of the study revealed that, although components of what was required existed in many systems, no system combined these components. Furthermore, few systems were designed to be used in classrooms in which apparatus manipulation and substantial pupil-pupil interaction were the predominant surface features.

Most of the recent major surveys of classroom processes (e.g. Soar, 1973; Brophy and Evertson, 1974; Stallings and Kaskowitz, 1974; Solomon, 1976) have opted for the use of multiple category systems rather than engaging in instrument development. This seems to be in line with the general feeling that a saturation point has been reached in such instruments and that new efforts, in the absence of major theoretical breakthroughs, simply yield redundant results. Nevertheless, because the purposes of this study were somewhat more narrow than those of the major surveys, and since more emphasis was required on the substantive content of classroom behavior (an aspect which is relatively neglected in classroom observation research), it was decided that some effort would be devoted to constructing an instrument that met the specific requirements of the study.

The Bellack System. One of the more well known (if not widely used) classroom analysis systems, that was developed by Bellack and his associates (Bellack, *et*. *al*., 1966) had a basic structure that was compatible with the primary requirement that both substantive and instructional dimensions of behavior be observed. The Bellack system is also one of the few in existence that has a clearly articulated theoretical base (at least with respect to the basic unit of analysis) that permits the investigator to go beyond the overt surface structure of the observed events. The first part of the study was therefore devoted to developing a modified version of the Bellack system that would meet the remaining requirements of the study.

The theoretical basis of the Bellack system derives from Wittgenstein's concept of language games (Wittgenstein, 1953), a concept which seems to have been first introduced in the context of classroom observation research by Aschner (1961). This concept led Bellack to search for the basic "moves" of the ongoing language game. Four such "pedagogical moves" were identified;

structuring (STR), soliciting (SOL), responding (RES), and reacting (REA). These moves were used both to identify units of discourse (i.e. as coding units) and as the basis for the analysis of interactions (i.e. as units of analysis). Beyond this basic dimension, divisions of each move into substantive and instructional content were made, with each of these dimensions having further subdivisions (substantive-logical and instructional-logical) corresponding to what may be termed the cognitive level (e.g. factual, analytic, evaluative, etc.) of the specific substantive or instructional content. Within the instructional dimension a separate "extra logical" set of categories was identified corresponding to an individual's "rating" of a previous move (e.g. the teacher might rate a student's response positively or negatively and would express the rating by some verbal convention). The latter set of categories is especially interesting for this study because it helps operationalize the concept of control.

For purposes of this study, the pedagogical move was retained as one dimension of analysis and the distinction between the substantive and instructional dimensions was retained. Beyond this, the system underwent considerable modification. In addition, the modification process was used as a means of training a number of persons in the use of the system, particularly in the coding of transcripts. A brief description of the modified system follows. A more complete outline, including examples of coding, is found in Appendix A.

Recording and Transcribing. The procedures by which videotape recordings of classroom events were gathered have already been described. For coding purposes, transcripts were prepared of the audio portion of the recordings. These transcripts were then edited by a member of the project staff who had been present during the original recording. Timing marks, at six second intervals, were placed on the transcripts to facilitate the coding of physical actions when no verbal discourse was occurring.

Coding Unit. The basic coding unit was a single sentence (this represents a departure from the Bellack system in which the pedagogical move was used as the coding unit). Physical actions were coded separately by reference to the videotapes, using the transcripts as guides. If verbal discourse was occurring, the coding unit for physical action was congruent with that for the corresponding verbal component. If no verbal discourse was occurring, physical action was coded every six seconds. The time unit was converted to sentence equivalents for purposes of analysis. The latter procedure involved the use of the ratio of sentences utterred to total lesson time for each lesson and allowed the direct addition of physical action units coded with and without verbal discourse.

Lesson Phase. Each lesson was divided into five major phases using the following categories:

1. Problem Specification
2. Data Collection
3. Data Interpretation
4. Summary
5. Cleanup

Within each lesson phase, behaviors were considered to be either on task, indirectly on task, or off task, with appropriate codes assigned to these categories.

Speaker. Both the speaker and the person to whom the utterances was addressed were coded. Separate codes were used for teacher, pupil's partner (as receiver), pupil outside the group under analysis, and receiver indefinite.

Pedagogical Move. The four basic moves; structuring, soliciting, responding, and reacting were taken directly from the Bellack system, with minor modifications to allow for the direct empirical character of the lessons at hand. That is, references to observations, measurements, physical actions, and so on were common in science activities, but did not occur in the lessons used in the original development of the system.

Substantive-Logical. In the Bellack system, the analysis of the substantive dimension involved the use of two sets of categories, the first to designate the subject matter under discussion, and the second to refer to the cognitive processes involved. As it happens, in this study, cognitive processes, at least insofar as they involved data collection and interpretation, could be considered as the content of the lessons. There was no particular interest, for example, in electricity as subject matter in a lesson using batteries and bulbs. Rather the concern was with teaching controlled experimentation, graphing and other aspects of the process of science. In the modified system, therefore, substantive (subject matter) categories were disregarded and the substantive-logical (cognitive process) categories modified to reflect the empirical character of the lessons. For this purpose, the science process categories developed by the American Association for the Advancement of Science for the program Science - A process Approach (AAAS, 1968) were used. In addition, the substantive-logical processes used by Bellack were retained. Although there might appear to be some overlap between, say, observing and fact stating, it was found that a useful distinction could be made between references to the immediate and empirical problem at hand and more general discussion of subject matter if the two sets of categories were used.

Originally, a "level of thought" was assigned to each substantive-logical category, using a three point rating scale. The coder was thus required to make judgments on whether the speaker was demonstrating low, medium, or high level of ability to use the particular process coded. Later, it was found that instances of low and high levels were so rare (either because of unreliability of coding or because such instances did not, in fact, occur) that the levels concept was judged to serve no useful purpose. While this presented a major problem in terms of the original objectives of the study, alternative procedures for assessing levels of ability in the processes were devised for the experimental study so that it was not necessary to rely on the coding system for this purpose.

Process Categories

Observing
Classifying
Quantifying
Communicating
Predicting
Inferring
Hypothesizing
Defining Operationally
Interpreting Data
Controlling Variables
Experimenting
Formulating Models

Product Categories

Fact Stating
Defining
Describing
Interpreting
Explaining
Evaluating
Opining
Justifying

Control. For purposes of this study, the "instructional meaning" dimension of the Bellack system was reconceptualized in terms of the notion of control. While the distribution of pedagogical moves between teacher and pupils gives some indication of degree of teacher control, a more detailed analysis can be made. A distinction was first made between declarative, interrogative, and imperative sentences. Further subcategories were then used to indicate more specifically the nature of the controlling act.

Sentence Type	Controlling Categories
	Teacher Requirement
	Pupil Requirement
	Performance
	Procedure
Statement	Statement of Intention
Request →	Attention
Command	Repetition
	Assistance
	Clarification
	Elaboration
	Example

Type of Reference. In terms of the concept of control as the basic function of verbal discourse in the classroom, the instructional meaning dimension of the Bellack system may be interpreted as identifying a set of referents for the controlling function. A modified version of Bellack's instructional categories was thus used to classify the types of referents that occur. Again, the categories are designed to reflect the empirical character of the science activities.

Statement
Logic
Language Mechanics
Assignment
Procedure
Person
Action General
Action Physical
Action Vocal
Action Cognitive
Action Emotional
Apparatus
Individual Results
Class Results
Expected or Desired Outcome
Observed Phenomena
Recording Equipment or Personnel
External Objects, Events, etc.

Ratings. In Bellack's study an attempt was made to assign "instructional logical" categories to utterances of an instructional nature in the same way as substantive-logical meanings were assigned to substantive utterances. For this study, however, these categories seemed to contribute little to the concept of control and were therefore not included. The exception was for those categories labelled "rating reactions". These were moves which served to evaluate previous statements. Since such ratings had an obvious bearing on the control question, the rating reaction idea has been retained as a separate dimension of the current system. The categories have been modified from Bellack as follows:

Positive
Qualifying
Repeating
Negative
Positive/Negative (request for a rating only)

Physical Actions. An obvious surface feature of the classes used in this study was the large amount of physical activity associated with apparatus manipulation and data collection and recording. It was the desire to capture the essence of this activity that led to the decision to use videotapes in recording the lessons. The types of physical actions of interest were those that seemed to have the potential to shed light on teacher-pupil and pupil-pupil relationships and particularly on pupil ability to perform the required tasks in collecting and interpreting data. In the absence of a physical action dimension in the Bellack system, a set of categories was generated with reference to the videotaped data available.

Apparatus Manipulation

Setting Up
Dismantling
Adjusting
Taking Reading or
  Measurement
Moulding, Shaping,
  Cutting, etc.
Watching
Nonspecific Action
No Action

Recording

Writing
Drawing
Graphing
Tabulating
Calculating
Nonspecific Recording

Management Movement

Fetching/Returning
Changing Position
  (moves to another group)
Raising Hand
Off-Task
Cleanup
Recording Equipment
Nonspecific Management
Not Codable

Coder Training and Reliability. The coding of transcripts was carried out by a team of four coders, all of whom had been directly involved in developing the system. It was thus possible for coders to attain increased proficiency in coding while the system was under development. This also tended to make coders sensitive to such problems as confusion in operationsl definitions or categories, omissions in successive drafts of important behaviors, and the particular "sets" of other coders. The process also made it possible to approach the problem of reliability of coding from a developmental perspective.

At an early stage in the process, transcripts were coded by a pair of coders working together, developing a consensus on each sentence as they worked through a transcript. However, analysis of transcripts coded by separate teams indicated that members of a team tended to develop "sets" which led to systematic differences between teams. Thus, a single error, say, in interpreting the difference between two similar categories would be repeated many times, leading to a reliability estimate that was much smaller than expected from random errors alone.

In an attempt to identify the sources of systematic error, as well as to measure the degree of random error, a second stage coding procedure was devised, whereby two coders independently coded each transcript. A matrix comparison was then developed in which the frequency of agreement and disagreement on each category could be determined. A sample matrix for pedagogical moves is given in Table 2. The diagonal elements of this matrix indicate the frequency of agreements on each category, while the off-diagonal elements give the frequencies of disagreements between specific pairs of categories. As can be seen, most disagreements occur between a relatively small number of pairs of categories. The elimination of these disagreements could significantly improve the reliability of codings. The computer program which generated the comparison matrices also provided a list of sentences on which large significant disagreements occurred. This list was used as the basis for discussion among all coders of the sources of error.

Originally, it was intended to use as the final coding, the corrected codes after the discussion among coders was complete. However, it was found that these discussions, if continued long enough, would lead to the elimination of almost all disagreements. Spuriously high reliabilities could thus be obtained, apparently simply through the process of wearing down dissenters within the group. It was thus necessary, before final coding took place, to devise a third procedure for the detection of systematic errors without creating the illusion of near perfect reliability through consensus.

The final procedure used was one in which each transcript was coded by a single coder and this coding reviewed by a second coder.

Table 2

Sample Comparison Matrix: Pedagogical Moves

		Coder A					
		Not Coded	STR	SOL	RES	REA	NOC
Coder B	Not Coded	4	4	0	0	0	0
	STR	0	29	2	3	2	0
	SOL	0	12	188	17	4	1
	RES	0	5	6	66	8	0
	REA	0	28	12	21	67	1
	NOC	0	14	11	7	6	6

At this stage, coders were reasonably well aware of each other's biases and could watch for instances of systematic error. These errors could then be resolved through discussion.

The above procedures make it difficult to report a set of reliability coefficients because indices based on comparisons between individual coders would be depressed due to the presence of systematic errors which could be adjusted. On the other hand, indices based on the final degree of agreement would likely be artifically high due to the collaboration between coders. As a compromise, a set of coefficients of agreement based on the number of sentences changed by a reviewer is presented in Table 3. It should be noted that although the percentages reported were calculated before the adjustment of errors by coder discussion, these percentages were obtained after substantial experience in the review procedure. At this stage, it is possible that coders were sufficiently aware of each other's biases and conventions to substantially reduce the error rate. The percentages should be therefore considered somewhat higher than might be obtained through other coding techniques.

Table 3

Sample Percentages of Agreement Between Coder and Reviewer

Dimension	Percent Agreement Transcript A 955 Sentences	 Transcript B 805 Sentences
Speaker	100	100
Receiver	100	99
Task	100	100
Move	90.6	85.7
Substantive-Logical	94.8	90.1
Give-Request-Command	92.8	89.5
Controlling	92.8	87.7
Reference	90.0	88.4
Rating	86.3	72.0

V TEACHER AND PUPIL BEHAVIORS: AN OVERVIEW

The results to be presented in this chapter are based on the application of the category system to the thirteen lessons recorded in the first phase of the study. These data serve to provide an overview of the range of teacher and pupil behaviors and teacher-pupil interactions that occur in the type of lesson under investigation. The analysis is designed particularly to shed light on the possible variations in teacher control that may occur in the natural (non-experimental) classroom setting, thereby providing a preliminary picture of the degree of variation that might be imposed under experimental conditions while maintaining the representativeness of the experimental setting with respect to normal classroom conditions. The experimental treatments applied in the second phase may, in this sense, be considered as attempts to define and apply the extremes of teacher control that occur under normal classroom conditions. The analysis also gives some indication of the variation in pupil behavior to be expected and of the nature of teacher and pupil roles in this type of classroom setting.

Because of sampling limitations, the results presented here must be considered highly tentative and the conclusions drawn therefore somewhat speculative. First, the results are applicable only to science lessons having process goals and characterized by active pupil participation in data gathering and interpretation. More important, the generalizability of statements about teacher behavior is limited because no data are available on the stability of teacher behaviors across occasions. Thus, what appears here as differences between teachers may conceivably simply be instances of day to day differences within teachers or of differences between classes. Nevertheless, as statements about possible variations that occur naturally in such classrooms, the teacher behaviors described are useful in spite of generalizability problems. The pupil sample is less restricted in size and representativeness than the teacher sample. Thus statements about variations in pupil behavior are likely to be more precise than those about teacher behavior.

The analysis is divided into three sections. The first deals with teacher behaviors, with emphasis on variables that seem indicative of variations in control. Next, a description of pupil behaviors is given, including comparisons by group, class, and grade level where differences between pupils at these levels exist. Finally, a more specific feature of the lessons, that of teacher interventions in the work of pupil-pairs, is discussed. Besides shedding further light on the issue of control, this analysis provides a basis for comparing teacher-pupil relationships in a small group setting with those in the whole class.

The analysis is primarily descriptive in nature, consisting mainly of mean proportions for different categories of behavior and the ranges of these proportions. Proportions are used rather than frequencies in order to adjust for differences in lesson lengths. Although the range is rarely used as a measure of variability because of its sensitivity to extremes, it is this sensitivity which makes the range a particularly useful statistic in this context since it gives some indication of the extremes towards which teacher behavior might be directed in the experimental study.

Teacher Behaviors

Teacher behaviors could be coded for two specific conditions within the lessons at hand; on occasions in which the teacher worked with the class as a whole (hereafter called teacher-class interactions, even though some of these interactions may be with individual pupils), and whenever the teacher happened to participate in the activity of one of the pupil groups being recorded (hereafter called teacher interventions). This section discusses teacher-class interactions while teacher interventions are dealt with in a later section.

In keeping with the aim of exploring, through these data, the question of teacher control, a number of general indicators are available of the degree to which the teacher is the dominant individual during teacher-class interactions and of the variations between teachers that occurred in the sample classes. Three such indices are reported in Table 4. The proportion of lesson time spent on problem specification (PRB) and data interpretation (DAI) indicate the degree of detail in which teacher instructions on the problem and method were given and the degree to which the teacher wished to pursue a common goal in the use of the data. It is, of course, possible that within these lesson phases the teacher allowed substantial pupil input in terms of defining the problem and interpreting the results. The total proportion of teacher talk is thus given as an indicator of teacher dominance within these lesson phases.

One point that is immediately apparent from these data is that the amount of teacher talk conforms to what has commonly come to be expected in conventional classes. This suggests, in a preliminary way, what will become more obvious as the discussion proceeds, that the degree of "openness" that seems to exist on the surface in these lessons may not conform to the realities of the lessons revealed through the more quantitative analysis.

The second point worth making is, of course, the wide variability in teacher behaviors which, it should be noted, is not due to a single extreme teacher but is characteristic of the data. That is, the ranges reported are more broadly representative of the data than is usually expected from this statistic.

Table 4

Some Preliminary Indicators of Teacher Control of Lessons

Variable	Mean	Range
Proportion of Sentences in Problem Specification Phase	.12	.03 - .26
Proportion of Sentences in Data Interpretation Phase	.22	.08 - .53
Proportion of Teacher Talk in Teacher-Class Interactions	.64	.47 - .87

Table 5 gives the mean proportion for each pedagogical move engaged in by the teachers and the range of variation of these moves. Two features are notable from this table. First, the amount of variation within each move is somewhat less than the total variation in teacher talk. This suggests that teachers are rather more consistent in their patterns of moves than in their total talk. Second, like total talk, the pattern of moves shows no significant departures from what would be expected in a more conventional lesson.

This point is illustrated more clearly by the comparisons in Table 6. In spite of differences in context, Bellack's original generalization about the teacher as primarily a solicitor seems to hold. Again, the inference is that teacher-class interactions have a surprisingly stable structure even in supposedly "open" classes. The increase in structuring moves in the present context is no doubt a function of the increased amount of general teacher "direction giving" associated with the use of apparatus. The reason for the decrease in reacting moves is less clear but is likely associated with the more concrete nature of the discourse in elementary science classes. Teachers may find it less necessary to verbally rate pupil utterances when these utterances refer to apparatus, observed phenomena and the like than in situations in which pupil utterances are more abstract.

Table 5

Means and Ranges for Proportions of

Each Pedagogical Move in Teacher Talk

Move	Mean	Range
Structuring	.23	.08 - .33
Soliciting	.53	.44 - .63
Responding	.03	0 - .09
Reacting	.18	.09 - .26

Table 6

Comparison of Proportions of Total Teacher Talk for Each Pedagogical Move for Four Studies

		Move			
Study	Context	Structuring	Soliciting	Responding	Reacting
Bellack (1966)	15 High School Social Studies Classes, United States	.03	.47	.06	.40
Lundgren (1972)[a]	8 High School Math. Classes, Sweden	.16	.40	.05	.34
Power (1973)	4 Eighth Grade Science Classes, Australia	.08	.50	.03	.38
Present Study	13 Sixth Grade Science Classes, Canada	.23	.53	.03	.18

Note. Proportions do not add to 100 percent due to rounding errors.

[a] A fifth move, labelled "individual help" (HEP) accounts for .05 of teacher talk in the Lundgren Study.

Some indication of the tone of teacher control can be obtained from the proportions of declarative interrogative and imperative sentences by teachers and the proportions of the various types of rating reactions that occurred. Means and ranges for these categories appear in Tables 7 and 8. The relatively large proportion of questions seems surprising since the lesson format would suggest that the teacher would have to spend some time giving directions on how the activity is to be carried out. An examination of the transcripts, however, revealed that teachers were, in fact, engaging in the classical recitation technique during the initial phase of the lesson, in that suggestions were obtained from pupils on how to proceed. The teacher than tended to rate these suggestions in such a way that the correct procedures were eventually made clear. The second major source of questions consisted of teachers calling on pupils to report results to the class. Again, this points to the use of conventional lesson procedures. The relatively small proportion of commands indicates that teachers do not rely strongly on imperatives in exerting control.

The distribution of ratings is characterized by more extreme variability than that for other categories. It seems clear that positive ratings exceed negative ones considerably. However, it is the use of the qualifying and repeating ratings that is potentially most interesting since these are more suggestive of teacher follow-up than the positive and negative ratings. Apparently some teachers

Table 7

Proportions of Declarative, Interrogative and Imperative Sentences in Teacher Talk

Sentence Form	Mean	Range
No G/R	.19	.10 - .29
Declarative	.28	.16 - .40
Interrogative	.38	.27 - .50
Imperative	.15	.09 - .19

Table 8

Proportions of Rating Reactions

By Teachers

Type of Rating	Mean	Range
Positive	.32	.04 - .45
Qualifying	.16	.03 - .31
Repeating	.27	.03 - .58
Negative	.17	0 - .31

commonly repeat pupil responses while others rarely do so. Whether this is simply a form of acknowledgement of a mechanism to ensure that the whole class benefits from the response is not clear. The variability in qualifying suggests that teachers differ in their tendency to make use of a pupil response, since the qualifying category suggests teacher modification of a response.

Further insight into the nature of teacher control is provided by the categories within the controlling dimension. Proportions of teacher discourse in these categories are given in Table 9. The consistently high proportions of sentences in the Performance category is worthy of note because sentences coded in this category were those in which the teacher either commanded or requested a specific immediate act on the part of an individual or the group. What emerges here is a view of the teacher as the director of activity - a picture quite consistent with the conventional teacher role. The fact that Performance statements were more prevalent than Procedure, Clarification or Suggestion categories indicates that the emphasis was on requiring immediate rather than longer term action on the part of students.

The specific contexts in which teacher activity occurred is revealed by the proportions for categories in the Type of Reference dimension as given in Table 10. While the contexts of the statements are more varied than the controlling categories, it is

Table 9

Proportion of Controlling Statements

By Teachers

Category	Mean	Range
Teacher Requirement	.01	0 - .04
Performance	.49	.41 - .61
Procedure	.14	.07 - .28
Intention	.01	0 - .06
Attention	.02	0 - .06
Repetition	.02	.01 - .05
Clarification	.09	.02 - .21
Suggestion	.01	0 - .04

Table 10

Proportions of Type of Reference Statements

By Teachers

Category	Mean	Range
Statement	.01	0 - .03
Procedure	.05	0 - .16
Person	.01	0 - .03
Action General	.05	0 - .09
Action Physical	.03	0 - .15
Action Vocal	.05	0 - .16
Action Cognitive	.06	0 - .23
Apparatus	.39	.19 - .63
Individual Results	.11	0 - .26
Class Results	.10	0 - .24
Expected Outcomes	.03	0 - .11
Phenomena	.02	0 - .05
Recording Equipment	.01	0 - .02
External Events	.03	0 - .16
Other or No Reference	.05	

notable that much of the discourse centers around apparatus. Although this type of discourse is obviously unique to lessons in which apparatus is used, and thus reveals little about how these lessons contrast with more conventional lessons, the strong emphasis on apparatus is further evidence of the immediate manipulative nature of teacher behavior. That is, teacher behavior seemed directed towards getting students to carry out certain immediate manipulative acts using the apparatus.

A second characteristic of the Type of Reference data is of interest. References to class and individual results seem more prevalent and more variable than other types of references (except apparatus). This shows that the practice whereby the teacher calls on pupils to report findings either individually or as a group is a common type of behavior but is used differently by different teachers. A check on whether particular teachers tended to use one form of request (individual or group) for reporting more commonly than the other revealed no clear pattern. Some teachers were high on both forms, others low on both forms, thus suggesting that one form of reporting is not used in preference to the other.

The final dimension of teacher behavior to be examined is that of the substantive-logical. Because of differences in the science processes emphasized at the different grade levels, systematic grade level differences were expected in this dimension. The summary given in Table 11 is thus broken down by grade level.

The first notable feature of these data is that more than half the teachers utterances showed no S/L content. This, combined with the fact that almost all utterances showed instructional content of some form, indicates that the emphasis in the lessons was more strongly on procedural matters than on the processes which formed the lesson content. The variations between classes within a grade are more prominant than the between-grade differences, although the latter does tend to show the expected directions, with decreased emphasis on observing and classifying and increased emphasis on quantifying for grade six.

Pupil Behaviors and Pupil-Pupil Interactions

The description of teacher behaviors presented in the foregoing section was designed specifically to indicate the potential variation in teacher control that might occur in the type of class setting under investigation. The rationale for the description of pupil behaviors, to be presented here, is more diffuse. What is being aimed at is an overview of pupil behavior which sheds light on questions such as pupil dominance, degree of emphasis of substantive versus procedural matters, whether some pupils adopt a "teacher" role in the group setting, the influence of matching versus mismatching of groups based on pupil characteristics, and whether substantive behaviors can be used as a means of judging the process abilities of pupils. Also of interest are the questions

Table 11

Proportions of Teacher Sentences in Substantive-Logical Categories By Grade

Category	Grade 2		Grade 4		Grade 6	
	Mean	Range	Mean	Range	Mean	Range
No S/L	.53	.37 - .73	.58	.41 - .66	.55	.44 - .74
Observing	.08	0 - .20	.11	.01 - .18	.02	0 - .05
Classifying	.13	0 - .28	.11	.01 - .34	<.01	0 - .01
Quantifying	.08	0 - .46	.00	.00 - .00	.17	.12 - .23
Communicating	<.01	.00 - .01	.06	.04 - .08	.04	.01 - .05
Predicting	<.01	.00 - .02	.01	.00 - .03	.01	.00 - .04
Inferring	.05	.00 - .17	.04	.01 - .09	.03	.01 - .07
Interpreting Data	.04	.00 - .18	.06	.00 - .16	.08	.04 - .15
Other	.08		.03		.09	

of group isolation, or whether pupils operate primarily within the defined group structure, and of group dependence on the teacher. These latter questions can yield some insight into the question of teacher control as it is manifested in pupil behavior in the absence of the teacher.

Again, the approach taken is descriptive. Mean proportions and ranges of these proportions for various coded categories are given and these proportions interpreted in terms of the questions of interest. The approach here is, however, somewhat more comparative than for the discussion of teacher behavior because the pupil sample size more readily permits comparisons, because data on pupil characteristics and group structures are available, and because a representative (although small) sample of pupils for each class is available. Grade level comparisons are particularly interesting for pupil behaviors because of age and developmental level differences in pupils in the three grades. Again the findings must be regarded as suggestive rather than as definitive. The emphasis is on hypothesis generation rather than on hypothesis testing. For this reason, the question of statistical significance is not emphasized in the comparisons to be made.

In interpreting the data on pupil behavior, it should be noted that virtually all pupil behaviors occurred in the context of pupil-pupil interaction. The proportion of sentences uttered by an individual pupil during teacher-class interactions was generally negligible (although when summed over all pupils, the proportion was fairly substantial). More sentences were utterred during teacher interventions in group work. However, this aspect of pupil behavior is treated separately in the analysis.

Table 12 gives the proportions of the total sentences in a lesson contributed by individual pupils. Considering the pupil pair as the basic unit for recording, it can be seen that, overall, a pupil pair contributed just over half the total discourse in a lesson. The remainder would be accounted for by teacher utterances and by the contributions of other pupils (primarily during teacher-class interactions but occasionally by the presence of other pupils engaging in discourse with the pair being recorded).

The range of variation in pupil utterances is perhaps of more significance than the mean. The overall range of .05 to .54 reflects a wide scatter between the limits rather than isolated extremes. In an attempt to determine whether this variation was a function of individual, group, or class differences, appropriate comparisons were made. These comparisons yielded the following results.

1. The ratio of participation for members of a pair ranged from about 1:1 to about 5:1. In six of the 25 groups, one pupil talked about twice as much or more as his partner. Considerable differences between pupils are thus indicated.

Table 12

Proportions of Sentences Uttered by Individual Pupils By Receiver and Grade

Grade		Proportion of Total Sentences in Lesson	Proportion of Pupil Sentences Received By		
			Teacher	Partner	Other
2	Mean	.31	.20	.78	.02
	Range	.10 - .54	.03 - .46	.50 - .95	.00 - .16
4	Mean	.22	.19	.72	.09
	Range	.05 - .53	.04 - .42	.46 - .91	0 - .47
6	Mean	.29	.11	.86	.03
	Range	.11 - .44	.02 - .30	.51 - .98	.00 - .19
Overall	Mean	.28	.18	.79	.04
	Range	.05 - .54	.02 - .46	.46 - .98	.00 - .47

2. Differences between groups within a class tended to be relatively small. In only one case did the number of sentences uttered by one group exceed twice that of the second group within a class. Even this variation, however, is surprising considering that the opportunity for pupil-pupil interaction was the same for all groups within a lesson since lesson length and total teacher talk were the same within a lesson.

3. Between class differences tended to be large. Pupils in the groups analyzed contributed from .13 to .82 of the total discourse. This suggests the possibility that differences between pupils can be accounted for by differences between classes. However, the within-group differences referred to in (1) suggests that pupil differences occur that are independent of class differences.

In summary, differences between pupils within a group can be interpreted in terms of pupil dominance, while differences between groups within a class is clearly related to opportunity for pupil discourse, which itself is a function of lesson format (i.e. of teacher control).

As Table 12 indicates, most pupil talk was directed towards that pupil's partner, with a mean proportion of .79, compared to .18 for the teacher and only .04 for other pupils. The latter proportion is especially interesting because it indicates that the groups remained relatively isolated from each other. This reinforces the argument that the teacher does maintain relatively high control over group activities since, in the absence of such control (either explicitly or implicitly maintained), much more inter-group interaction might be expected. As was the case for proportions of total talk, substantial between-pupil, between-group, and between-class differences could be found. For example, in nine of the 25 groups, one pupil talked to the teacher proportionally more than twice as much as his partner. Differences in pupil-to-partner interaction are more difficult to interpret because of the reciprocal nature of this type of interaction and because the proportions are dependent on the proportions directed towards the teacher, as well as the total number of sentences uttered by an individual.

The proportions for pedagogical moves, as given in Table 13, appear to show a greater balance for pupils than for the teacher. However, these figures reflect differences in setting that make comparisons difficult. In the case of pupils, most moves reflect interaction with the partner and are therefore not independent of each other. That is, for example, a large number of solicits by one member of the pair would tend to result in a large number of responses by his partner. Also, the pupil-pupil situation is clearly one in which the participants are, at least nominally, in the same role.

Table 13

Proportions of Each Pedagogical Move for Pupils

Move	Mean	Range
Structuring	.13	.01 - .31
Soliciting	.31	.05 - .56
Responding	.14	.05 - .29
Reacting	.33	.21 - .59

The salient features of the data are 1) the relatively high proportion of reacts, 2) the imbalance of solicit and respond moves, and 3) the wide range within each move, particularly for solicits. The first probably reflects a coding convention whereby all references to direct observations were coded as reacts. Many more such references occurred for pupil-pupil interaction than for teacher-class interaction. The imbalance in solicits and responds reveals an interesting feature of pupil-pupil discourse in that many pupil solicits either required a nonverbal response (e.g. a physical action involving apparatus manipulation) and were therefore not coded. The difference between the proportion for solicits and that for responses is, in fact, a reasonable measure of the amount of nonverbal responses. The wide variation between individuals is a further indicator of differences in roles assumed by members of a group. In fact, when the dominant pupil in a group, in terms of total talk, was identified (e.g. a pupil who talked more than twice as much as his partner), in five out of six groups that pupil was also dominant in terms of number of solicits. The wide range within each move might be expected to reflect the reciprocal nature of the moves. However, this was found not to be the case. That is, there was no particular tendency, say, for extremely high solicits by one pupil to be reciprocated by extremely high responses by his partner.

It should be noted that no systematic grade, group, or class level differences were found to occur for the proportions of moves.

The proportions of pupil discourse involving substantive-logical content averaged less than half for all grades. The major pattern, in which observing is predominant for grades 2 and 4 and quantifying is predominant for grade 6 is similar to that for teachers. However, the proportions for classifying in grades 2 and 4 and for interpreting data in grade 6 are smaller than the corresponding proportions in the teacher data. This indicates an important aspect of teacher and pupil roles whereby pupil-pupil discourse was largely on data collecting processes, while discourse involving the teacher was more concerned with drawing some form of generalizations from the data. Thus, grade 2 and 4 pupils might, for example, observe the characteristics of objects while the teacher would be more concerned with the grouping of objects into classes. Similarly, grade 6 pupils would carry out many measurements. The generalizations resulting from these measurements would then tend to be developed by the teacher.

The proportions for sentence forms, given in Table 15 reveal an interesting feature of pupil-pupil discourse in that commands exceed requests by a ratio of more than 2 to 1. This is in strong contrast to the proportions for teachers, in which the ratio is approximately reversed (the actual proportions in both categories is

Table 14

Proportions of Pupil Sentences in Substantive-Logical Categories By Grade

	Grade 2		Grade 4		Grade 6	
Category	Mean	Range	Mean	Range	Mean	Range
No S/L	.77	.60 - .90	.58	.40 - .81	.66	.49 - .79
Observing	.10	.02 - .24	.28	.10 - .39	.04	0 - .10
Classifying	.02	0 - .08	.03	0 - .13	-	-
Quantifying	.02	0 - .29	-	-	.20	.08 - .31
Communicating	-	-	.04	0 - .14	.05	.01 - .12
Predicting	.02	0 - .04	-	-	-	-
Inferring	.01	0 - .05	.02	0 - .05	.01	0 - .04
Interpreting Data	-	-	.01	0 - .03	.02	0 - .09
Other	.02	-	.04	-	.02	-

Table 15

Proportions of Declarative, Interrogative and Imperative Sentences for Pupils

Sentence Form	Mean	Range
No G/R	.42	.26 - .58
Declarative	.29	.16 - .50
Interrogative	.09	0 - .23
Imperative	.21	.03 - .33

larger for teachers, reflecting the greater proportion of soliciting moves). Examination of the transcripts revealed that pupils often seemed to engage in "ordering each other around", especially with respect to apparatus manipulation. Pupils used many expressions of the form "no, don't do that", "move it a little more", or "add some more weight" and so on. In contrast, pupils seemed rarely to question each other, except to ask for a measurement or an observation. Thus pupil use of questions is substantially different from teacher use, again reflecting obvious role differences.

It is instructive to examine the issue of pupil dominance from the standpoint of proportion of commands. Using the extreme groups in terms of total pupil talk (one pupil talked more than twice as much as his partner), it was found that the dominant pupil in five of the six such groups was also dominant in terms of proportion of commands issued. This corresponds precisely with the results for solicits as compared to total talk, suggesting that some consistency exists in the pattern of discourse of dominant pupils. It remains for future studies to determine whether relationship exists between group dominance and outcomes.

Table 16 reveals that, within the controlling dimension, the preponderance of statements was for the performance category, although this proportion for pupils is smaller than that for teachers, reflecting the smaller proportion of solicit moves.

Table 16

Proportions of Pupil Sentences in Controlling Categories

Category	Mean	Range
Teacher Requirement	.01	0 - .03
Performance	.31	.19 - .46
Procedure	.02	0 - .09
Intention	.05	0 - .16
Attention	.05	0 - .17
Repetition	.06	.01 - .15
Clarification	.06	0 - .23
Other or no Controlling	.44	-

Also, in contrast to teacher behavior, the proportion of sentences coded in the procedure category is much smaller. This indicates that pupils were less concerned with overall lesson procedures, with giving directions for conduct of the activity as a whole, and generally with any process beyond the immediate manipulation of apparatus or collection of data.

Pupil statements, like teacher statements, referred to apparatus more than to any other category. However, pronounced grade level differences occurred. Balancing the reduced emphasis on apparatus in grade 4 was an increased proportion of references to phenomena. Again, these differences clearly were related to differences in lesson content. The grade 4 activities required pupils often to make a particular change in materials and to observe the consequences. For example, pupils were required to drop food coloring in water or to mix different types of liquids. Reporting of the observed behavior of the materials thus became an important aspect of the grade 4 work.

The final dimension of pupil behavior to be considered is that of ratings. The proportions for the various rating categories are shown in Table 18. Here, the pattern for pupils is strikingly different from that for teachers. The preponderance of negative ratings, especially in grades 2 and 4 is particularly notable. A systematic shift with grade level appears to exist for both positive and negative ratings, with a tendency to become more positive and

Table 17

Proportions of Type of Reference Statements for Pupils By Grade

Category	Grade 2		Grade 4		Grade 6	
	Mean	Range	Mean	Range	Mean	Range
Statement	-		-	-	-	-
Procedure	-		.03	0 - .09	.02	0 - .07
Person	.02	0 - .05	.02	0 - .08	.02	.01 - .04
Action General	.02	0 - .09	.03	0 - .06	.03	.01 - .09
Action Physical	.02	0 - .09	.09	.02 - .26	.05	.01 - .08
Action Vocal	.03	0 - .08	.02	0 - .07	.02	0 - .04
Action Cognitive	.02	0 - .07	.02	0 - .08	.02	0 - .09
Action Emotional	.05	.01 - .14	.01	0 - .02	.02	.01 - .06
Apparatus	.43	.28 - .55	.24	0 - .48	.50	.39 - .59
Individual Results	.03	0 - .07	.09	0 - .18	.04	0 - .12
Expected Outcomes	.02	0 - .15	-	-	-	-
Phenomena	.09	.03 - .19	.25	.13 - .38	.07	.01 - .10
Recording Equipment	.05	0 - .16	.02	0 - .09	.02	0 - .04
External Events	.01	0 - .03	-	-	-	-
Other or no Reference	.21	-	.18	-	.19	-

Table 18

Proportions of Rating Reactions for

Pupils By Grade

Type of Rating	Grade 2		Grade 4		Grade 6	
	Mean	Range	Mean	Range	Mean	Range
Positive	.23	0 - .50	.31	0 - .80	.41	.11 - .63
Qualifying	.02	0 - .07	.05	0 - .29	.08	0 - .21
Repeating	.05	0 - .17	.05	0 - .25	.05	0 - .16
Negative	.69	.43 - 1.00	.58	.20 - 1.00	.44	.17 - .89

less negative with increasing grade level being manifested. The source of the large proportions of negative ratings, though not of the grade level differences, is revealed by closer examination of the transcripts, where it is found that pupils tended to react to each other's actions with admonitions of the form "no, don't do that", "that's wrong, dummy", and similar expressions. The use of such expressions is related to the frequency of commands in that many of the negative expressions were in the imperative form. It should be noted that, although the use of positive ratings by pupils was as high as that for teachers, ratings in the reinforcement sense, such as those used by the teachers for motivation or encouragement, were largely absent in pupil discourse. Similarly, the relatively small proportions of qualifying and repeating ratings for pupils, as compared to teachers, reflect teacher and pupil role differences. The tendency for the teacher to repeat a pupil statement for the benefit of the whole class, or as a neutral rating device, is, of course, absent in discourse between pupils. Similarly, pupils rarely found occasion to modify each other's statements as teachers would tend to do when probing for correct responses.

It should be noted that the high degree of instability in the rating categories is essentially a function of the relatively small number of sentences (less than 10 percent of the total) which involved ratings. These wide ranges thus reflect extremes resulting from the calculation of proportions from small frequencies. Because of this, the conclusions regarding the comparison of proportions across categories and between pupils and teachers should be treated with caution. Nevertheless, the differences are of such magnitude and are generally sufficiently consistent to justify the conclusion that the pattern of use of ratings is a useful description of pupil behavior and of teacher and pupil roles.

Physical actions of pupils perhaps deserve a brief comment. Only the "no action" and the "apparatus adjustment" categories showed frequent occurrence (.32 and .26 respectively). Variations in these categories showed stronger between-class than between-pupil or between-group effects, suggesting that lesson structure was the determining factor in the nature of physical action. The category "watching" also showed striking inter-class variations, although the average proportion for this category was quite small. Again, lesson structure, particularly teacher involvement, seemed to be the important underlying variable.

Teacher Interventions

The description of teacher-group interactions to be presented here gives a somewhat unique perspective on the structure of lessons since it reveals the teacher in a situation of having to devote attention to only two pupils and it gives these pupils a unique opportunity to interact with the teacher free from the influence of a large number of other pupils. The specific rationale for the study of teacher interventions stems from a possible contrast between teacher-pupil relationships in the class and group settings. The nature of the science activities under investigation requires a departure from the "teacher front and center" arrangement of more conventional classes. On the surface, it would appear that this would naturally lead to a reduction in teacher dominance of classroom activities. In fact, it seems clear that a reduction in teacher dominance was one of the aims of curriculum developers in opting for a mode of classroom operation based on small group work. A study of teacher-pupil relationships in the small group setting should therefore shed light on the degree to which this aim is being achieved in the classroom.

Some of the specific questions addressed in the following discussion are as follows: Do interventions occur on the basis of pupil demand or does the teacher appear to retain control of the allocation of her time? Do some pupils have greater access to teacher time than others? If so, is this a matter of pupil demand or of teacher choice? What are some of the reasons for the initiation of interventions? How do patterns of pedagogical moves during interventions compare to those in other settings? How long and how frequent are the interventions for a particular group? What changes in nature of discourse occur from beginning to end of an intervention. The findings which bear on these questions are interpreted in the light of the overall issue of teacher control as it relates to the larger study.

The analysis follows the general pattern established in the discussion of teacher and pupil behaviors in that the various dimensions of the category system are used. In addition, however, data on numbers of interventions, the length of interventions, and more general interpretations of the processes by which interventions are initiated, continued, and terminated are given. Also, because of the issue of possible shifts in teacher and pupil behavior as a function of the change in setting, direct comparisons are usually made with the teacher and pupil behaviors already described.

An intervention was defined as occurring whenever a sentence with the teacher as speaker and one of the pupils being recorded as receiver appeared embedded in a series of sentences involving the pair of pupils being recorded, or if one of these pupils requested teacher attention and received a response. Interventions were identified through a manual search of the coded data. A total of 149 interventions were identified in the ten lessons used.[4]

Table 19 summarizes the findings for intervention numbers and lengths. These findings suggest that, in general, teachers spent a relatively short time with a group on each occasion, although some notable exceptions to this trend did occur. Some teacher

Table 19

Summary of Findings for Intervention Length and Number

Variable or Comparison	Result
Intervention Length	Range 1-63 sentences, mean 7.4 S.D.3.6. Distribution highly skewed with sharp drop in frequency with increased length. Isolated instances only of interventions longer than 15 sentences.
Number of Interventions Per Group	Range 3-16, mean 7.9, S.D.3.9. Distribution approximately rectangular.
Grade Level Comparisons	No significant differences by ANOVA. Instances of very long interventions confined to sixth grade, likely reflecting increased lesson length.
Group and Class Comparisons	No significant differences between groups within a class; significant between class differences for grade 2 only.

Table 19 (continued)

Variable or Comparison	Result
Correlation of Length and Number of Interventions	R = −.26, p<.05
Pace of Discourse	Discourse more rapidly paced during intervention than when pupils worked without teacher.

attention was also given to all groups. This seems to suggest that the teacher maintained a fairly high degree of control over time allocation and that, with few exceptions, particular groups did not command an inordinate amount of teacher time (the major exception of the 63 sentence intervention occurred in a rather lengthy grade six lesson). The negative correlation between intervention length and intervention number (with intervention number adjusted for variations in lesson length) also supports this argument.

The following is a brief summary of findings related to the initiation of interventions and the sequence of events once the intervention commenced:

1. Teachers and pupils accounted for almost exactly an equal number of initiations (75 for pupils, 74 for teachers).

2. For short (1-3 sentences or one teaching cycle) teacher initiated interventions, requests outnumbered statements as initiating moves 39 to 25 (the absence of commands as initiators is notable). Declarative openings were characterized by statements about procedure, which pupils generally accepted without comment. Among the requests, progress reports (18), and requests for clarification (9) occurred most often, with pupils generally complying with the request.

3. For short student initiated interventions, the situation was reversed, with declarative openings occurring more than twice as often as requests. Giving progress reports and reporting observations were the most common declarative openings, while requests were either for clarification or for the teacher to give procedural statements.
4. The pattern for initiation of longer interventions closely paralleled that for short interventions. As the discussion proceeded, pupil reporting of progress and observations, and pupil requesting and teacher giving clarifications remained predominant. However, a dramatic shift occurred near the end of these longer interventions. These almost invariably ended with the teacher making a statement about procedure.

Although the fact that the pupils initiated about half the interventions seems to suggest a relatively high degree of pupil control of teacher time, this result must be interpreted in the light of the fact that each class consisted of some 15 groups and that experience indicates that simultaneous requests for teacher attention by several groups is common in such classes. Under these conditions the teacher must often ignore some requests while attending to others. While no direct evidence is available on how the teacher selects particular groups for attention it does seem surprising that the teacher can initiate half the interventions under the conditions described. This suggests that the teacher does retain substantial control over the allocation of her time among groups. A similar conclusion follows from the fact that none of the observed groups totally lacked teacher attention.

Further evidence for this conclusion lies in the tendency for teachers to break off longer interventions by giving directions about procedure. This presumably serves to ensure that the group can continue the activity in the teacher's absence and permits the teacher to give attention to other groups.

The proportion of teacher and pupil sentences uttered during interventions compared to those during teacher-class interactions are shown in Table 20. Considering the large difference in the teacher-pupil ratio in the two settings, these proportions seem remarkably stable. Nevertheless, it should be noted that the proportions are as notable for their inter-class variation as for the similarity of means. Since the correlation of .47 between teacher talk during and outside interventions does not reach significance at the .05 level, it is not possible to argue that proportion of teacher talk is related to the teacher rather than to other variables in the setting.

Table 20

Proportions of Teacher and Pupil Talk During Interventions and During Teacher-Class Interactions

Proportion		Interventions	Teacher-Class Interactions
Teacher	Mean	.64	.56
	Range	.47 - .87	.40 - .69
Pupil	Mean	.36	.44
	Range	.13 - .53	.28 - .60

It is difficult to judge the significance of these findings in terms of expected changes from class to group setting. On the one hand, the reduction in the number of individuals in the group setting should allow more opportunity for a single individual (including the teacher) to talk. At the same time, the reduced formality of the setting and the consequent reduced need for the teacher to orchestrate the setting should allow for more pupil talk. In fact, of course, individual pupils do talk considerably more in the intervention setting. What is remarkable is the relative stability of mean teacher talk in the face of an apparent large change in the setting. Individual pupil talk seems to change drastically to meet the requirements that the teacher behave similarly towards the group and the class.

Further information on the stability of behaviors in intervention and non-intervention settings can be found in the comparison of pedagogical moves in the two settings, which appear in Table 21. Again, in spite of the change of setting, the pattern of teacher moves appears quite stable. This is particularly true for soliciting, the primary teacher move. (The ranges for moves were also substantially narrower than for total talk. These ranges are omitted from the table for brevity). A comparison of teacher and pupil moves in the intervention setting reveals the pattern of teacher-solicit/pupil response that has become familiar in studies using the Bellack system. It is important to

Table 21

Mean Proportions of Teacher and Pupil Discourse for Each Pedagogical Move

	Teacher		Pupil	
Move	Non-Intervention	Intervention	Non-Intervention	Intervention
STR	.23	.15	.13	.05
SOL	.53	.56	.31	.32
RES	.03	.10	.14	.41
REA	.18	.19	.33	.03

note, however, that the proportion of pupil solicits was somewhat higher and that for responses somewhat lower than those found in most studies. For example, in the original Bellack study (Bellack, _et_. _al_., 1966), the proportion of all pupil moves in the solicit category was only .13 while the proportion of responses was .69. Similarly, Lundgren (1972) found proportions of .12 and .60 for pupil solicits and responses respectively. The Power (1973) study revealed an even more extreme situation, with some 92 percent of all pupil discourse being responses and only 5 percent solicits. The small group setting thus seems to be associated with a substantial change in pupil behavior from both the whole class setting and the pupil-pupil setting, even though, in the broadest sense, the pattern of teacher solicit-pupil response was maintained.

Proportions of sentence types shown in Table 22 again confirm the relative stability of the pattern of teacher utterances, and the contrast with that for pupils. Since requests and commands were both associated with the solicit move, the Table 22 data serves to refine the argument, in that it is clear that most teacher solicits were in the form of requests. The earlier data on initiations suggests that these requests were commonly for progress reports or for clarifications. The proportion of pupil commands seems incongruous with teacher and pupil roles. However, as already indicated, in the non-intervention setting most pupil commands were directed at the other pupil and tended to refer to

Table 22

Proportions of Statements, Requests, and Commands for Teachers and Pupils

	Teacher		Pupil	
Category	Non-Intervention	Intervention	Non-Intervention	Intervention
Statement	.28	.28	.29	.40
Request	.38	.35	.08	.13
Command	.15	.20	.21	.16
Not Coded	.19	.17	.41	.31

apparatus manipulations or to observations. It seems reasonable to argue that this also accounts for the commands during the interventions.

Tables 23 and 24 give the proportions of sentences in each of the categories in the type of reference and rating dimension. The relative stability of teacher behavior is again the most salient feature of these data. In the case of references, teacher and pupil proportions follow similar patterns. This might be expected, since the flow of discourse tends to be such that continguous sentences are likely to be related to each other in content, as reflected in these categories. In contrast, teacher and pupil use of ratings differ sharply. Pupils show much greater use of negative ratings but rarely qualify or repeat. The latter seems to be a clear instance of different pupil and teacher roles, since the repeating or modifying a pupil response seems to be a common teacher device. It is interesting to note that pupils become much more positive and much less negative during interventions. Again, this appears to reflect the presence of the teacher. In the pupil-pupil setting it was quite common for a pupil to admonish his partner (using statements of the type "no, don't do that!"), especially with reference to apparatus manipulation. In the intervention setting, it would be highly unlikely that the pupil would react to the teacher in this manner, even though he might continue to do so towards his partner.

Table 23

Proportions of Type of Reference Categories for Teachers and Pupils

Category	Teacher		Pupil	
	Non-Intervention	Intervention	Non-Intervention	Intervention
Procedure	.05	.08	.02	.06
Actions (Physical, Cognitive, Vocal)	.10	.18	.15	.09
Apparatus	.39	.50	.40	.38
Individual Results	.11	.09	.05	.13
Class Results	.10	–	–	–
Expected Outcomes	.03	.03	.01	.03
Phenomena	.02	.03	.12	.07
TOTAL[a]	.89	.91	.75	.76

[a] Remaining proportion divided among other categories or not coded in this dimension.

Table 24

Proportions of Ratings for Teachers and Pupils

	Teacher		Pupil	
Category	Non-Intervention	Intervention	Non-Intervention	Intervention
Positive	.32	.38	.30	.50
Qualifying	.16	.13	.04	.04
Repeating	.27	.24	.05	.04
Negative	.17	.25	.61	.42

Summary and Discussion

This chapter presented descriptive data on teacher and pupil behaviors under three variations of the overall class setting; teacher-class interactions, pupil-pupil interactions in the absence of the teacher, and teacher interventions in group activities.

With respect to teacher behaviors, the findings indicated a remarkable stability of behaviors across the two variations in which the teacher was involved. These behaviors also conformed, in general, to teacher behaviors observed in other studies using the same categories. An observed increase in the proportion of structuring moves in comparison to the studies could be attributed to context differences between classes observed in this study and those used in other studies. Specifically, the "hands on" nature of the science activities required that the teacher spend considerable time giving general directions for apparatus usage, data collection, and so on. This type of behavior would be absent from more conventional classes. A similar argument can be advanced to account for the decreased proportion of teacher reacting moves.

There is an indication from the data that much classroom discourse is concerned with immediate procedural matters rather than with substantive matters. This is true for both teachers and pupils. At the same time, substantive behaviors did show the expected grade level variations, based on differences in lesson content.

The fact that pupil behaviors, particularly pedagogical moves, changed substantially from the non-intervention to the intervention setting, even though the surface change was simply from a two-person group to a three-person group, clearly indicates that the addition of the teacher leads to a substantial change in person roles in the group. Interventions more closely resemble teacher-class interactions than they do pupil-pupil interactions.

In terms of the behavior setting construct, the results suggest that the setting may be better defined by teacher and pupil roles than by surface features such as group size. Similarly, it would appear that a shift from strong to weak framing, in the Bernstein sense, must involve more than a change in grouping structure or a move to place pupils in an active manipulative role rather than in the role of attending to the teacher as found in conventional classes.

A further notable feature of pupil behavior, which has implications for teacher control, was the relative isolation of the pupil groups from each other. Individual pupil utterances were almost invariably directed towards either the teacher or the partner. In view of the extensive opportunity for contacts outside the group offered by the relatively low level of teacher involvement with each group, this is a surprising finding. It does suggest that pupils, for whatever reason, do tend to adhere to roles defined by the basic structure of the lesson (once groups are formed by the teacher, these groups become the dominant aspect of communication within the classroom).

Two of the findings for pupil behavior in the group setting, the relatively high proportion of imperative sentences and the high use of negative ratings, at first appeared puzzling. However, examination of the transcripts showed that pupils tended to issue orders with respect to apparatus manipulation and to admonish each other when something would go wrong (or be perceived as wrong by one of the pupils). It would be interesting to see whether such behavior characterize interactions between children in small groups outside the classroom setting (e.g. children at play) or whether the behaviors are classroom-specific.

Some evidence existed of a tendency in some groups for one of the pupils to be dominant. This was shown by one pupil consistently engaging in more overall talk, more soliciting moves, and more commands than his partner. This, of course, raised the question of whether such behavior was related to pupil characteristics.

VI EXPERIMENTAL RESULTS

The findings of the descriptive study suggest that substantial variation between teachers can occur even within the constraints imposed by the nature of the curriculum and the general lesson format derived from that curriculum. This is true even though, on the average, teacher behaviors and teacher and pupil roles are not drastically different in this context from those found in more conventional class settings. It is possible, of course, that the change in surface characteristics of a lesson from whole class to small group work and from lecture, recitation or demonstration to laboratory activities leads to increased *variability* in teacher behavior. In any case, the question of variations in teacher behavior can have some influence on classroom processes and outcomes, independently of curricular influences and of other constraints external to the teacher.

The purpose of the experimental study, as already indicated, was to investigate systematically the question of variations in teacher behavior, specifically in relation to what has been termed "teacher control". The theoretical and practical arguments for this type of study have been developed in Chapter 2 and will not be repeated here. Suffice it to say that the study was designed to test two sets of hypotheses, one pertaining to treatment main effects and the second to interactions of treatments and pupil characteristics. These hypotheses were investigated with respect to several types of dependent variables including pupil achievement, preference, perceptions of lessons, and classroom behaviors.

To recapitulate the essential elements of the experimental design, ten sixth grade classes were exposed to two nine-week treatments, defined in terms of a number of teacher behaviors derived from the classroom observation system. All classes were given both treatments, in alternating sequence. Classes were taught by four teachers trained in the curriculum content and in the implementation of the treatments. Teachers were randomly assigned to classes, with each teacher having at least one class in each treatment. Teachers were assigned to new classes at the end of the first nine-week round in order to minimize unique teacher-class effects.

This chapter presents the findings of the experimental study. First, data on the implementation of treatments is given. This is followed by the results for achievement, preference, pupil perceptions, and classroom behaviors. Emphasis throughout is placed on findings that replicate across the two treatment rounds and that are comparable across similar criterion variables. This avoids the problem of attaching too much importance to individual tests of statistical significance in a situation in which many such tests were carried out.

Implementation of Treatments

Table 25 gives the variables on which teachers were expected to differ in implementing the high and low control treatments. In general, the differences are in the desired direction. Exceptions occur for teacher responding and reacting, which show essentially no differences and pupil reporting individual results, which shows a difference in the opposite direction from that predicted. In the latter case, however, the large inter-lesson variance within a treatment renders the result unreliable. In any case, this form of pupil behavior was perhaps not under teacher control to a sufficient degree to show predictable effects. The results for the soliciting move, while in the desired direction, showed a smaller difference than was anticipated. This seems to reflect the prominance of this move in defining teacher role and its relative insensitivity to changes in setting, as already indicated in the comparison of teacher-class and teacher-group interactions in the previous chapter.

Because the findings for certain of the dependent variables suggested that teacher differences and treatment-teacher interactions tended to be of some importance, it was decided to develop a breakdown of certain teacher behaviors relevant to the treatments by teacher and treatment. Results of this breakdown appear in Table 26. It is apparent from these data that the order of magnitude of teacher differences is at least as great as that for

Table 25

Comparison of Behaviors in High and Low Control Treatments

Variable	High Control	Low Control
Teacher-Class Interaction	876	419
Teacher-Group Interaction	671	1,095
Proportion of Pupil Talk During Teacher-Class Interaction	.06	.10
Proportion of Pupil Talk During Teacher-Group Interaction	.24	.29
Teacher Structuring	453	369
Teacher Soliciting	539	650
Teacher Responding	89	84
Teacher Reacting	229	227
Ratio of Commands to Total Solicits	.69	,79
Pupils Reporting Individual Results	76	46
Teacher Substantive-Logical Discourse	187	230
Pupil Substantive-Logical Discourse	34	55

Table 25 (continued)

Variable	High Control	Low Control
Average Wait Time (Sec)	1.5	1.9

Note: Units are total sentences uttered in a sample of approximately 130 minutes lesson time for each treatment. Sample includes 13 lessons with all teachers included at least once in each treatment. Ratios and wait time computed for same sample data.

treatment differences. In spite of the specific nature of the training and the uniformity of lesson content, it appears that teachers did behave in some ways characteristic of the individual.

Achievement

It will be recalled that the method of analysis for the achievement criterion was a generalized multiple linear regression analysis, using the individual pupil as the unit of analysis, with variables entered in the following order: (a) initial ability (Raven and CTBS composite entered simultaneously); (b) teachers; (c) treatment-teacher interactions; (d) ATI's with each attribute variable entered in stepwise order; and (e) treatments.

Table 27 presents the results of the regression analysis for the two rounds. As is usually the case, the ability measures accounted for the largest proportion of variance. The treatment effect was nonsignificant in both cases. Interaction effects accounted for six and eight percent of the variance in Rounds 1 and 2 respectively. In each case, the largest single interaction was with an attitude measure; attitude to school in Round 1 and attitude to science in Round 2. Since these two measures were relatively highly correlated ($r = .53$), the effect of these in a single regression equation is likely to depend on the order in which each was entered. That is, once the component of variance due to one of these variables had been extracted, the second would

Table 26

Treatment by Teacher Comparisons for Selected Teacher Behaviors

Variable	Treatment	Teacher 1	Teacher 2	Teacher 3	Teacher 4
Teacher-Class	HC	12	26	18	30
Interaction	LC	9	8	15*	11
Teacher-Group	HC	26	17	7	7
Interaction	LC	26	27	13*	29
Teacher	HC	11	10	8	12
Structuring	LC	8	9	15*	7
Teacher	HC	11	14	11	15
Soliciting	LC	15	15	6*	19
Ratio of	HC	.36	.29	.40	.27
Commands to	LC	.21	.24	.54*	.17
Total Solicits					
Teacher	HC	4	11	4	2
Reacting	LC	3	8	8*	5

Note: Except for Ratio of Commands to Total Solicits, units are mean number of sentences utterred during a three minute segment.

*Round 2 data only.

Table 27

Summary of Regression Analysis, Science Process Achievement

Source of[a] Variance	Round 1[b] ΔR^2	Round 1[b] F	Round 2[c] ΔR^2	Round 2[c] F
Ability (IQ, CTBS)	.364	46.12***	.500	114.34***
Teachers	.007	.89	.037	6.02*
Treatment x Teachers	.005	.63	.040	7.03**
x Attitude to School	.038	10.11**	.002	1.11
x Attitude to Science	<.001	---	.010	5.40*
x Sex	.011	2.97	.003	1.66
x Locus of Control	.002	.54	.008	4.50*
x Neuroticism	<.001	---	.011	6.08*
x Extraversion	<.001	---	.004	2.12
Treatment	.004	1.02	<.001	---

*P < .05, **P < .01, ***P < .001

[a] Interactions omitted from Table were nonsignificant in either round.

[b] N = 164

[c] N = 232

be expected to make no substantial additional contribution. It seems reasonable to conclude that the results across the two rounds represent a stable interaction of attitude and treatment. Except for the attitude effect, the remaining interactions were not replicated over the two treatment rounds. The fact that all other interactions occurred for the second round suggests the possibility that the treatments may have been more clearly differentiated in the second round. However, since the teacher behavior data were not tabulated by round, this cannot be directly established.

Preference Scale

Findings for the various main and interaction effects for the preference scale appear in Table 28. On the surface, it would seem that a relatively large number of significant effects emerged for this measure. However, application of the criterion of consistency across replications indicate that only the treatment teacher-teacher and treatment-locus of control effects are stable. The absence of a significant treatment effect in the second round is notable, in view of the highly significant first round result and the fact that the mean treatment difference in the second round was in the same direction as but was substantially greater than that for the first round (2.29 versus .76 points).

Table 28

Summary of Regression Analysis,

Preference Scale

Source of[a] Variance	Round 1[b] ΔR^2	Round 1[b] F	Round 2[c] ΔR^2	Round 2[c] F
Teachers	.026	2.15	.036	2.85
Treatment x Teachers	.140	13.35***	.116	10.26**
x Sex	.060	12.25***	<.001	---
x Locus of Control	.053	11.91***	.014	4.28*
x CTBS	.007	1.47	.085	24.95***
x Dependency	.008	1.85	.022	6.62*
x Attitude to School	.019	4.36*	.002	.60
x Self-Concept	.024	5.02*	<.001	---
x Dominance	.007	1.62	<.001	---
x IQ	.005	1.16	<.001	---
Treatment	.050	12.31***	.002	.61

*P < .05, **P < .01, ***P < .001

[a] Interactions omitted from Table were nonsignificant in either round.

[b] N = 164

[c] N = 232

In any case, as Table 29 reveals, the treatment-teacher interaction seems to be of considerably greater importance than the treatment main effect, especially since the result for Teacher 1 replicates quite well over the two rounds. Since Teacher 1 taught different classes in the two rounds, it seems reasonable to attribute the treatment-teacher interaction to some aspect of that teacher's behavior rather than to class effects. This is not revealed, however, by the teacher comparisons presented earlier since there is nothing in these comparisons to suggest that Teacher 1 was unique.

Choice Task

The distributions of pupil choice of high and low control tasks are shown in Table 30. It is evident that pupils demonstrated a strong tendency to choose low control. A cross-tabulation of choice by treatment with pupils had just completed revealed that the choices were essentially independent of the immediate treatments. It is interesting that a fairly large number of pupils·changed their choice in favor of low control at the end of the second round. It is possible that this may be due to pupils having had the benefit of both treatments and were thus more clear on the alternatives. However, this cannot be

Table 29

Cell Means for Treatment-Teacher Interactions, Preference Scale

Treatment	Teacher				Mean
	1	2	3[a]	4	
			Round 1		
HC	20.54	22.32	---	23.33	22.00
LC	24.87	20.47	---	22.40	22.76
Mean	22.65	21.48	---	22.87	
			Round 2		
HC	18.05	21.90	22.95	22.22	21.11
LC	24.59	22.14	23.17	23.60	23.40
Mean	21.27	22.02	23.06	23.00	

[a] Teacher 3 omitted from Round 1 analysis because no data were obtained for this teacher's low control class.

Table 30

Frequencies of Pupil Choice of High and Low Control Tasks

Choice	Observed	Expected	X^2
	Round 1		
Chose High	138	163	
Chose Low	188	163	7.67**
Total	326	326	
	Round 2		
Chose High	90	162	
Chose Low	234	162	64.00***
Total	324	324	

**P <.01, ***P <.001

directly confirmed from the data.

A summary of findings for the cross-tabulations of pupil choice with pupil characteristics and with teachers appears in Table 31. Nonindependence of choice and a particular pupil characteristic, as determined by the chi-square test, can be interpreted as evidence of interaction.

It can be seen that the result for teachers seems to parallel that for the preference scale. However, the contingency table did not present such a clear picture as did the table of means for the preference scale, in that the tendency to choose low control was not uniquely associated with pupils taught by teacher 1.

Strong sex and dominance effects and a moderately strong dependency effect appeared for the choice task. The sex and dependency effects are especially interesting since each appeared in one of the rounds for the preference scale. There thus seems to be reason to argue that these are important effects since they replicate to that degree over two similar criteria.

It should be noted, of course, that the method of analysis for the choice task was less powerful than for the preference scale. In particular, to the extent that student characteristic variables are correlated, then the interactions reported may not be independent of each other. This likely accounts for the presence of both a dominance and a dependency effect in the choice task since these two variables are significantly correlated ($r = -.44$)

Table 31

Summary of Chi-Square Tests for Pupil Choice of High or Low Control Task by Pupil Characteristics

	Round 1			Round 2		
Variable	χ^2	DF	P	χ^2	DF	P
Teacher	9.68	3	<.05	8.60	3	<.05
Sex	12.86	1	<.001	15.77	1	<.001
CTBS Composite	3.81	2	NS	4.16	2	NS
SES	3.50	2	NS	2.88	2	NS
Extraversion	4.07	2	NS	3.23	2	NS
Neuroticism	3.98	2	NS	2.88	2	NS
Dominance	15.35	2	<.001	19.21	2	<.001
Dependency	6.17	2	<.05	7.12	2	<.05
Locus of Control	2.04	2	NS	5.43	2	<.10
Self Concept	4.44	2	NS[a]	9.67	2	<.01
Attitude to School	2.22	2	NS	1.68	2	NS
Attitude to Science	2.14	2	NS	1.65	2	NS

Interview Schedule

Because the tabulation of individual interview items over four rounds yielded a large number of contingency tables, the results are presented in summary form in Table 32. This summary is based on a clustering of items having the following common themes:

1. Pupil perceptions of lesson in content, process, procedural or other terms.
2. Teacher structuring of lesson.
3. Teacher versus pupil role.
4. Adequacy of time.
5. Pupil-partner relationships.

For items in cluster 1, there was evidence of response differences that were related to a number of the independent variables of interest. A summary of the chi-square tests for this cluster is thus presented in Table 33. The contingency tables from which these chi-square tests were computed consisted of item response, classed as content, process, procedure, or other crossed with the independent variable listed, either in dichotomous or trichotomous form.

This table indicates that pupil perceptions are strongly related to both treatments and teachers. In addition, dominance and extraversion were related to response in three of the four rounds.

Table 32

Summary of Pupil Responses to Interview Item Clusters Based on Four Rounds

Item Cluster	Results
1. Perception of Lesson in Content, Process, Procedural or other terms	Generally more pupils perceived the lessons in content terms, and relatively fewer in process terms. However, in round one there were more pupils in high control treatment who perceived the lessons in procedural or other terms. There were significant differences in all four rounds due to teachers in the way the pupils perceived the lessons.
2. Teacher Structuring	Generally all pupils, irrespective of treatments perceived the teacher as the source of procedure. Pupils about evenly divided on whether they could have carried out activity with less teacher direction.

Table 32 (continued)

Item Cluster	Results
3. Teacher versus Pupil role	Irrespective of treatments, in round one teacher role was perceived as more important than pupil role because of need to follow teacher directions and importance of obtaining results the teacher wanted. The trend changed over rounds two, three and four. In round four high control perceived teacher role more important, whereas low control pupils perceived pupil role as more important.
4. Time	Irrespective of treatment, pupils generally felt that there was adequate time to carry out the activity.
5. Pupil-Partner Relationships	Pupil pairs generally worked well together. Pupils tended to give partners more credit, or thought that both shared the work but very rarely gave credit to themselves for performance in the activity. No differences found due to treatment.

Table 33

Summary of Chi-Square Tests for Pupil Responses to Interview Items Cluster One for Rounds One to Four

		Round 1		Round 2		Round 3		Round 4	
Variable	Df	χ^2	P	χ^2	P	χ^2	P	χ^2	P
Treatment	3	24.24	<.001	5.02	N.S.	8.65	<.05	9.93	<.05
Teacher	9	36.74	<.001	39.29	<.001	28.16	<.001	23.81	<.01
Sex	3	14.72	<.01	5.63	N.S.	7.71	N.S.	10.52	<.05
SES	6	46.06	<.001	11.06	N.S.	11.23	N.S.	10.83	N.S.
IQ	6	16.81	<.001	10.61	N.S.	13.11	<.05	4.76	N.S.
Extraversion	6	13.55	<.05	22.79	<.001	5.29	N.S.	15.45	<.05
Neuroticism	6	12.42	N.S.	33.70	<.001	6.62	N.S.	6.25	N.S.
Dominance	6	17.14	<.01	28.72	<.001	12.58	<.05	10.02	N.S.
Dependency	6	7.15	N.S.	8.42	N.S.	6.90	N.S.	10.55	N.S.
Locus of Control	6	6.69	N.S.	14.24	<.05	15.96	<.05	7.15	N.S.
Self Concept	6	3.74	N.S.	13.90	<.05	10.39	N.S.	6.37	N.S.

Nevertheless, examination of the contingency tables yielded no specific pattern of response replicated over the four rounds. Since it is unlikely that the number of significant effects appearing in the table could have occurred by chance, it is necessary to conclude that responses must have been dependent on some unmeasured variable. The possibility that responses were a function of the content of specific lessons is one such explanation. It is also possible that responses were a function of some complex set of interactions that could not be revealed by the contingency tables.

Items on teacher structuring of the lesson and on teacher and pupil roles yielded responses that were more closely associated with pupils perceiving the lessons more in high control than in low control terms. The teacher was generally viewed as playing a strong leadership role, particularly in the first round. In the final round, however, a shift in perceptions occurred with respect to the amount of work done by the teacher as compared to the pupils. This reinforces the possibility, alluded to earlier, that by the end of the second treatment round either the treatments were being implemented more faithfully or the pupils were finally beginning to experience a change in role perceptions. This tendency, however, cannot be considered as being especially strong since pupils continued to perceive the teacher as the source of procedures and the judge of results and continued to feel that it was very important to carry out the activity the way the teacher wanted.

Pupil Classroom Behaviors

As described in Chapter III, pupil behaviors were recorded by videotaping three to four three-minute time samples for each of five pairs of pupils on three occasions for each class. The videotapes were then coded using the observation system developed in the first phase. In this case, however, the coding was done directly from the videotapes rather than from transcripts as was the case in the descriptive study. While this necessitated some reductions in the number of categories coded, this was not considered to be a serious problem because experience in the descriptive study had given some indication of which categories might be useful and which could be reasonably reliably coded.

The recording sequence was such that, for each pupil pair, a three minute segment near the beginning, middle, and end of each lesson was available. Because of the greater likelihood of teacher-class interaction near the beginning and end of the lesson and because the nature of this type of interaction varied for these two segments, it was decided initially to carry out a separate analysis for each segment. Under these conditions, each dependent variable consisted of the number of sentences, for a particular category of the coding system, utterred within a three minute segment by each pupil. Frequencies were transformed to proportions where necessary to permit more direct comparisons across individuals.

From the outset, it had been anticipated that a number of problems would exist in compiling and interpreting these data. First it was found in the descriptive study that some potentially interesting categories (such as references to teacher requirements or to expected outcomes) occurred so rarely that no confidence could be placed in their reliability. Second, there was no clear criterion for selecting potentially useful variables other than judgements that could be made on the basis of the descriptive study. The primary interest was originally in the substantive-logical and controlling dimensions; the first because it was argued that substantive-logical references would relate directly to achievement and to pupils' thought processes, and the second because variations in teacher control categories were expected to influence similar categories for pupils. In particular, interactions of pupil characteristics with treatments were expected. In view of the variations in teacher implementation of the treatments, as already described, there was also reason to expect treatment-teacher interactions. Finally, there was the impression that a great many other factors, including a pupil's relationship with his partner, the specific content of the lesson itself, as well as simple randomness in pupil behavior would contribute to the presence of a large error term in the analysis, thus reducing the chance of detecting systematic treatment or interaction effects.

The analysis followed the general form used for the treatment and preference scales. Regression equations were developed with each classroom behavior as a dependent variable, and with teacher, treatment-teacher interactions, ATI's, and finally treatment entered as predictors. One such equation existed for each variable, each three-minute segment, and each treatment round. Ten variables were selected for this analysis.

Because of the length of the tables, and the fact that very few significant effects emerged, the complete summary of the regression analysis is relegated to the Appendix. The findings may be briefly summarized as follows: (1) No treatment effects occurred for any of the variables, (2) Isolated ATI effect occurred but these were not sufficiently well replicated across segments and rounds to warrant further interpretation. It is likely that the few significant effects were instances of Type I errors, (3) Treatment-teacher interactions, replicated reasonably well across occasions, occurred for the variables of pupil no talk, and pupil listening to teacher, (4) Teacher main effects were also found for these variables. The latter two findings appear to relate to pupil opportunity for talk, which might be expected to relate both to treatment and in the manner in which the treatments were implemented by a particular teacher in a given lesson.

In order to pursue this point further, while reducing the complexity of the analysis, it was decided to attempt a supplementary analysis in which treatment, teacher, and treatment x teacher interaction effects were examined, while disregarding ATI effects and summing our lesson segments and treatment rounds. A unit of data for this analysis therefore consisted of the total number of sentences in each category utterred by a pupil over the three lesson segments during which that pupil was observed. Pupil characteristics, occasions, lesson length differences, and the like were considered to be randomly distributed throughout the cells of the model. Some confounding of class effects with the effects under analysis did exist. However, the combining of treatment rounds yielded data from at least two classes in each cell of the design.

The multiple regression analysis in this case was essentially equivalent to a two-way factorial analysis of variance. Results of this analysis appear in Table 34. As expected, in view of the large built-in error variance, the proportions of variance accounted for by the components of the model were relatively small. Nevertheless, a number of statistically significant effects did emerge. The treatment x teacher interaction was significant for five of the nine behavior variables, accounting for from about four to eight percent of the variance. Teacher affects accounted for from four to sixteen percent of the variance (in addition to the interaction variance) for the four variables yielding statistically significant differences. Treatment effects occurred for seven of the nine variables, accounting for from two to eight percent of the residual variance after the interaction and teacher effects had been extracted.

Because raw frequencies rather than ratios were used in this analysis, it seemed likely that the results for the various dependent variables might be highly correlated. In particular since the four pedagogical moves were mutually exclusive and additive, it was possible that the results for each move would simply mirror those for total sentences. The correlation matrix of Table 35 would seem to confirm that this is indeed the case. Table 36, however, suggests that the trends are more complex than this. For example, pupils taught by Teacher 1 talked about twice as much in low control as in high control. Yet, these pupils engaged in approximately the same number of soliciting moves in both treatments. Similarly, the amount of reporting of individual results for Teacher 1 shows a reverse effect from that of total sentences. There does seem to be a general tendency for the treatment main effect to be

Table 34

Summary of Regression Analysis,

Treatment x Teacher, Pupil Behavior Variables

Dependent Variable	Effect	R^2	ΔR^2	dF	F	P
Total Pupil Talk	Treatment x Teacher	.059	.059	3,277	4.74	<.01
	Teacher	.099	.040	3,224	3.31	<.05
	Treatment	.175	.076	1,223	20.54	<.001
Structuring	Treatment x Teacher	.105	.105	3,227	8.88	<.001
	Teacher	.172	.067	3,224	6.04	<.001
	Treatment	.213	.041	1,223	11.62	<.001
Soliciting	Treatment x Teacher	.025	.023	3,227	1.78	N.S.
	Teacher	.077	.052	3,224	4.21	<.01
	Treatment	.107	.040	1,223	9.99	<.01
Responding	Treatment x Teacher	.016	.016	3,227	1.23	N.S.
	Teacher	.042	.026	3,224	2.03	N.S.
	Treatment	.064	.022	1,223	5.24	<.05
Reacting	Treatment x Teacher	.044	.044	3,227	3.48	<.05
	Teacher	.058	.014	3,224	1.11	N.S.
	Treatment	.098	.040	1,223	9.89	<.01
Total Substantive/ Logical	Treatment x Teacher	.066	.066	3,227	5.34	<.01
	Teacher	.076	.010	3,224	.81	N.S.
	Treatment	.103	.027	1,223	6.83	<.01

Table 34 (continued)

Dependent Variable	Effect	R^2	ΔR^2	dF	F	P
Quantifying	Treatment x Teacher	.080	.080	3,227	6.58	<.001
	Teacher	.099	.019	3,224	1.57	N.S.
	Treatment	.118	.019	1,223	4.80	<.05
Apparatus	Treatment x Teacher	.029	.029	3,227	2.25	N.S.
	Teacher	.053	.024	3,224	1.90	N.S.
	Treatment	.066	.013	1,223	3.10	N.S.
Individual Results	Treatment x Teacher	.022	.022	3,227	1.65	N.S.
	Teacher	.185	.163	3,224	15.05	<.001
	Treatment	.192	.007	1,223	1.94	N.S.

Table 35

Correlations Between Pupil Behavior Variables

	PSENT	STR	SOL	RES	REA	TSL	QUA	APA	RET
PSENT		.84	.78	.48	.73	.30	.67	.55	.56
STR			.58	.34	.59	.05	.66	.55	.47
SOL				.23	.55	-.01	.41	.34	.72
RES					.23	-.01	.61	.50	.17
REA						-.03	.48	.40	.35
TSL							.02	.00	-.03
QUA								.78	.17
APA									.17
RET									

Table 36

Cell Means,[1] Treatment x Teacher,

Pupil Behavior Variables

Variable	Treatment	Teacher				Treatment Means
		1	2	3	4	
Total Sentences	1	10.63	7.00	6.40	8.45	8.45
	2	4.80	8.45	3.55	4.16	4.80
Structuring	1	3.93	1.45	2.07	2.35	2.41
	2	2.83	.48	.53	.20	1.11
Soliciting	1	2.50	1.23	1.07	2.00	1.74
	2	2.81	.97	1.03	.30	1.38
Responding	1	.90	1.08	.53	1.29	1.02
	2	.81	.51	.20	.55	.52
Reacting	1	1.80	.55	.47	1.55	1.13
	2	1.00	.21	.63	.05	.52
Total Substantive/Logical	1	1.43	2.73	2.20	2.00	2.13
	2	.94	1.37	1.77	.80	1.25
Quantifying	1	2.93	1.15	1.07	1.32	1.76
	2	1.29	.66	.17	.50	.67
Apparatus	1	1.56	.90	1.07	1.58	1.16
	2	1.13	.66	.13	.50	.62
Individual Results	1	1.47	.43	.67	1.00	.88
	2	4.39	1.00	.83	.00	1.72

---

[1] Units are sentences utterred in approximately nine minutes of lesson time.

consistently in the direction of greater frequencies for low control. Similarly, the teacher main effect seems to be such that Teacher 1 appeared to elicit more pupil talk than other teachers. It thus seems reasonable to conclude that main effects follow the general pattern established by the total sentence variable but that the interaction effect is somewhat more complicated. In any case, it is again evident that teachers tended to implement the treatments in different ways and that this prevents the drawing of straightforward conclusions about treatment effects.

VII SUMMARY AND CONCLUSIONS

Summary of Purposes and Methodology

This study was concerned with the general question of teacher control within the context of elementary science classes. The issue of control was of interest for both theoretical and practical reasons. Theoretically, Barker's behavior setting construct leads to the possibility that an individual's behavior may be a function of his perceptions of the goals of a setting and of his role in the setting. Similarly, Bernstein's distinction between strong and weak framing proposes that behavior may differ in accordance with the range of options open to the teacher versus the pupil in the setting. In more practical terms, the behaviors used to define teacher control were considered to be those which could be varied by the teacher given the usual constraints of curriculum, materials, and general lesson format. The specific interest in the elementary science context developed from this area and also from the specific structure of most of the major elementary science curriculums which require at least surface departure from conventional teacher and pupil roles towards what would appear to be less direct teacher control. Nevertheless, there was no evidence that the surface change (e.g. pupil use of apparatus, small group work) would result in a more fundamental change in roles or whether such a change would serve any useful purpose.

The first phase of the study was concerned with developing a classroom observation system capable of yielding data on teacher and pupil behaviors in both the whole class and the small group setting. It was also required that the system encompass both the substantive and instructional dimensions of behavior; the first because it was proposed that pupil substantive behavior within the classroom would yield a measure of pupil ability to deal with the scientific problems found in the lessons (e.g. would be, in some sense, a measure of achievement), and the second because this dimension subsumed the categories of teacher control that were of interest. This phase of the study was also concerned with developing an overview of teacher and pupil behaviors in regular classes in which science lessons were being conducted in the general mode described.

Videotapes of thirteen class sessions were used as the data base for the first phase. The starting point for the classroom observation system was the system developed by Bellack and his associates (Bellack, et. al., 1966). This system was modified to meet the requirements of the study, using the videotapes as a guide. Transcripts developed from the videotapes were then coded using the modified system. The coded categories became the base for the descriptive analysis of the class sessions presented in Chapter V.

The second phase of the study was designed to focus more closely on the question of teacher control and its impact on outcomes. Two experimental treatments were defined in terms of a number of categories derived from the classroom observation system. These treatments were intended to represent the poles of a continuum of control, within the limits that might be expected of teachers under ordinary classroom conditions given the constraints imposed by a specific curriculum, its associated materials, and a mandated general lesson format (pupils working in small groups each with its own apparatus). The study was conducted in ten sixth-grade science classes. Important elements of the experimental design included the use of both treatments in all classes in alternating sequence to permit a replicated analysis, the use of special teachers to allow teacher training and balanced assignment of teachers to classes, and the monitoring of treatments implementation through the use of videotapes and the coding of those teacher behaviors used to define the treatments. Dependent variables were achievement, pupil preferences, pupil perceptions of treatments as expressed in individual interviews, and pupil behaviors in the classroom based on categories of the classroom observation system. Data for the latter variables were derived from the coding of some thirty videotapes of class sessions, using a time sampling procedure to achieve a representative sample of pupil behaviors.

Summary of Descriptive Phase

A detailed description of the modifications made to the Bellack system to accommodate the purposes of this study has been given in

Chapter 4. Briefly, the major modifications were as follows:

1. The sender and receiver of a communication were both identified. In particular, individual pupils selected for recording, as well as the teacher, were identified.

2. The coding unit was changed from the pedagogical move to the single sentence to reduce ambiguities in deciding the boundaries of moves and to allow the coding of more than one substantive-logical or instructional category within a move. Pedagogical moves were, however, retained as a separate dimension for analysis.

3. The substantive dimension was modified to eliminate subject matter categories and to include, under the substantive-logical sub-dimension, science process categories derived from the process model used in the program Science-A Process Approach (AAAS, 1968). (The objectives of the lessons under analysis could be expressed in terms of teaching these processes).

4. The instructional dimension was revised and expanded to permit coding of statements, requests, and commands, and to include categories more directly representative of the empirical nature of the science lessons and more clearly emphasizing the focus on teacher control.

5. Categories for physical action were developed (although these categories proved to be less useful than originally anticipated).

Descriptive and comparative analyses of the data from the original thirteen class sessions were carried out for three settings within each lesson; the teacher-class setting, the pupil-pupil setting (a single pupil pair with the teacher absent) and the teacher intervention setting (a pupil pair with the teacher present). A detailed description of the findings is given in Chapter 4. These findings may be summarized as follows:

1. Teacher behaviors, averaged across teachers, showed a remarkable stability in the two variations of the class setting in which the teacher was involved. These behaviors also conformed, in general, to teacher behaviors observed in other studies in which the same categories were used, in spite of the fact that the surface features of the elementary science classes investigated here were substantially different from those of the more conventional (lecture-recitation) classes observed in other studies.

2. Substantial variation between teachers was found to occur, although these variations were not sufficiently great to negate the general pattern. For example, although the proportion of soliciting moves varied from .44 to .63 for

different teachers, there was no overlap between this range and the range for the structuring or reacting moves. that is, teachers always solicited more than they reacted or structured.

3. Teachers tended to ask questions more often than they gave commands or even made declarative statements. This held true for both teacher-class and teacher-group interactions.

4. When engaged in reactions to pupil statements, teachers tended to use more positive than negative ratings and tended to often qualify or, particularly, repeat pupil utterances.

5. More teacher utterances were related to immediate performance of some action on the part of pupils (reflecting the preponderance of soliciting moves). Statements about apparatus followed by those referring to individual or class results occurred most frequently.

6. Less than half the teacher utterances involved any form of substantive-logical content. Specific substantive-logical statements reflected differences in lesson content with grade level. Observing and classifying statements were predominant in grades 2 and 4 while quantifying was the most common category in grade 6.

7. Most pupil-pupil interaction took place within the specific pupil-pair assigned to work together. Pupil interactions with the teacher gave the next highest proportion, while interactions with pupils outside the assigned pairs were quite rare.

8. Differences in proportions of discourse between members of a group could be related to pupil dominance. Pupils who were dominant in terms of total discourse also tended to engage in more soliciting moves and to issue more commands.

9. Pupils tended to use many commands, especially in relation to apparatus manipulation. This was coupled with an extensive use of negative ratings, which suggested a picture of pupils "ordering each other around" through the use of negative reactions to each other's moves. Pupils, in fact, rarely questioned each other, preferring, it seems, to act and accept the partner's negative sanctions or else to give orders to the partner.

10. Pupil discourse, like teacher discourse, tended to be performance oriented, to reflect apparatus manipulation, and to show grade level differences in substantive-logical categories reflecting the differences in lesson content.

11. Pupil behaviors showed a marked shift when the teacher appeared to intervene in the activities of a group. Teacher interventions resembled teacher-class interactions more than small group interactions, indicating that teacher and pupil roles are quite strongly differentiated and are not susceptible to change with changes in the surface feature of the setting.

12. In general, teacher interventions were relatively short, task oriented, and appeared controlled by the teacher to a larger degree than expected in view of the competing demands for the teacher's attention that occur when a class is divided into small groups. Teachers more often initiated interventions with requests (e.g. for progress reports or for clarifications) than with declarative statements and almost never with commands. In general, pupils complied with the request that initiated an intervention.

13. Pupil initiated interventions were characterized more by declarative openings (primarily reporting observations or giving progress reports) than by requests. The latter type of opening usually called for a teacher procedural statement or a teacher clarification.

14. Longer interventions (more than 3 sentences in duration) followed similar opening moves to those for short inter-

ventions but showed a dramatic shift at the end, with the teacher almost invariably giving a procedural statement. The latter was interpreted as a teacher device for ensuring that behavior would be on task following the intervention or to permit the teacher to move to another group. These interpretations are consistent with the general picture of relatively high teacher control of the setting.

Summary of Experimental Findings

The detailed report of the findings of the experimental study appears in Chapter VII. The following is a brief outline of the major findings:

1. In general, teachers operated in a mode consistent with the operational definitions of the treatments. Nevertheless, differences between teachers were sufficiently great to warrant consideration of teacher effects in the analysis.

2. Results for the science process achievement test showed that pupil ability was the primary determinant of achievement, accounting for 36% and 50% of the variance in the first and second rounds respectively. Teacher effects and treatment-teacher interactions were statistically significant in the second round but failed to replicate over the two rounds. ATI effects accounted for only about 6% or the variance, with statistically significant effects occurring for pupil attitude variables. Treatment failed to account

for a significant proportion of variance, after all other effects had entered the regression equation. There was some indication that both teacher and ATI effects were more significant during the second round than during the first, suggesting the possibility that the treatments may have been more clearly differentiated, at least by certain teachers, in the second round.

3. Two criteria of pupil preference for treatments were used; a preference scale based on the variables used to define the treatments, and a choice task in which pupils were asked to carry out an activity in a chosen mode reflecting one or other of the treatments. Reasonably strong treatment effects, in the direction of preference for low control, were indicated on these criteria. Treatment-teacher interactions and ATI's were also statistically significant, although specific ATI effects did not replicate over the two criteria. The latter finding may be an artifact of the order of entry of variables in the regression equation, particularly since some of the pupil aptitude variables were intercorrelated.

4. The analysis of data from pupil interviews showed that relatively few pupils perceived the lessons in terms of their basic purpose, that of teaching certain aspects of

science as a process (specifically, the process skills of quantifying, interpreting data, and controlling variables). Pupils in the low control treatment tended to perceive lessons more in content terms, while the results for high control pupils, depending on the interview occasion, showed perceptions primarily as either content or procedure oriented. Pupils generally perceived the teacher as the source of information about procedure and as the judge of results. Early interviews showed that pupils perceived teacher role as more important than pupil role because of the need to follow directions and to obtain the results desired by the teacher. Later interviews showed treatment differences in perceptions. In particular, pupils in low control began to perceive their role as more important. Pupils generally worked well with their partners and tended to give their partners more or equal (but rarely less) credit for performance in the activity. Some ATI effects on perceptions were indicated but the nature of the interview data precluded a fully systematic investigation of these effects.

5. An analysis of pupil classroom behaviors, carried out on coded categories derived from samples of pupil behavior near the beginning, middle, and end of each lesson, revealed that treatment, teacher, and ATI effect accounted

for a relatively small proportion of the variance (of the order of 25 to 30 percent), depending on the variable under analysis. The exceptions were for teacher effects and treatment-teacher interactions for the variables of pupil no-talk (total proportion of pupil non-verbal behavior) and pupil listening to the teacher. These two variables were directly related to the manner in which teachers implemented the treatments in that the amount of teacher talk to the class as a whole (which would preclude pupil activity) was a function of treatment.

6. A supplementary analysis of pupil behaviors, in which ATI effects were deleted and lesson segments and treatment rounds combined, yielded a number of statistically significant treatment, teacher, and treatment x teacher interaction effects (again, however, the proportions of variance accounted for were relatively small). The general trend in treatment and teacher main effects was established by the total talk variable. Pupils generally talked about twice as much in low control as in high. Pupils taught by one of the teachers also tended to talk more than those taught by others, independently of treatment. Findings for pedagogical moves, substantive-logical statements, and other categories followed this same pattern. Interaction effects were, however, more complex, confirming the impres-

sion that the treatments were in fact implemented in different ways by different teachers.

Conclusions

The results of the descriptive study strongly support the general conclusion that teacher and pupil classroom roles possess an inherent stability that does not fundamentally change even under an apparent drastic change in the surface features of the class setting. More specifically, the change from a teacher-class setting, as in a conventional recitation lesson or as in the introductory and summary phases of the science activities used in this study, to a teacher-group setting in which the group consists of only two pupils yields no substantial changes in teacher behavior or the nature of teacher-pupil interactions. On the other hand, pupil behaviors are significantly different in a setting in which the teacher is absent (as when two pupils work together as a group).

It might be argued, of course, that the very existence of pupil-pupil interactions, which is brought about by the use of the small group format, itself constitutes a major change in classroom structure. Nevertheless, in such lessons one is still left with the impression of a pervasive teacher presence, as evidenced by the average of four teacher interventions per group in a lesson, the tendency for teachers to initiate about half of the interventions in spite of competing demands from various groups, the tendency for teachers to end longer interventions by procedural statements, and the high degree of maintenance of assigned group structures (i.e. the absence of cross-group interactions).

More specific to the issue of teacher control, the preponderance of soliciting moves, the performance orientation of most teacher moves, the use of qualifying and repeating ratings as well as positive and negative ratings, the distinct shift in the nature of pupil talk during teacher interventions, the absence of teacher responses, and the relatively common reporting of pupil results in both teacher-class setting and the intervention setting all point to the specific nature of teacher and pupil roles as those of high teacher control and pupil subordination. It is notable that no matter how dominant an individual pupil was in the pupil-pupil setting, in no case did a pupil contribute more than about half the talk in an intervention setting (the average for a pupil pair in such a setting was only .36).

The finding that pupils tended to use a high proportion of commands and of negative ratings is puzzling. Examination of the transcripts for clues to the context of such usage showed that pupils tended often to give directions about apparatus in an imperative tone (e.g. "add another washer", "drop the red coloring in"). and to admonish each other when something would go wrong (e.g. "that's too heavy, dummy", "you got it all wrong"). This is in strong contrast to the questioning or suggestive approach typically used by the teacher in such situations (e.g. "what should we do next?", "suppose we add another washer?"). It is obvious, of course, that this "imperative-negative" form of discourse is characteristic of pupil-pupil relationships and not of pupil-teacher relationships. It is almost inconceivable, for example, that the pupil would use a statement such as "that's too heavy, dummy" to the teacher and, equally important, highly unlikely that the teacher would use terms of that nature to a pupil, at least in the present context. It would be interesting to explore whether behaviors of this sort characterize interactions between children outside the classroom (e.g. children at play) or whether these behaviors are classroom specific.

One of the objectives of the first phase of the study could not be achieved because the science process categories in the substantive-logical dimension did not yield the desired measure of the thought processes of pupils as they confronted the problem posed in the lesson. The main difficulty was a methodological one; the inability to code levels for the substantive-logical categories. (It was for this reason that it was necessary to use pupil interviews in the experimental study). However, the problem was not purely a methodological one. Most pupil discourse was not substantive-logical but, rather, was procedural in nature. That is, pupils were more involved in specific apparatus manipulations and similar activities than in substantive matters. Even that portion of the discourse that could be coded in substantive-logical terms was largely confined to observing, particularly in grades two and four. At the level at which the class activities were conducted, it is difficult to see how an incorrect or a unique or unusual observation could have been made. The absence of such observations led to the failure to identify levels of skill in observing. Had the pupils engaged in a significant amount of classifying, inferring, or interpreting data, for example, it might well have been possible to detect the desired levels. Even in grade six, where such processes might have been expected to occur and indeed, were explicit objectives of the lessons, the absence of these processes in pupil discourse is notable.

The general picture that emerges, then, is one of procedural rather than substantive orientation in the lessons. Pupils and teachers alike seemed more concerned with following procedure than with substance. Exceptions to this did occur, of course, as evidenced by the proportion of .46 devoted to quantifying by one second grade teacher. Unfortunately, it was not possible to clearly establish whether these exceptions were characteristic of the teachers or their particular classes or whether they were isolated cases (these differences were probably not related to lesson content since this was controlled within grades).

The descriptive findings, of course, highlight the question of whether differences in degree of teacher control could be brought about and, if so, whether such differences would have a significant influence on pupil behaviors and outcomes. In addition, the question of whether teacher control differences might interact with pupil characteristics to yield differential effects on these dependent variables was also of both theoretical and practical interest, particularly in relation to the developing body of research on aptitude-treatment interactions. The experimental study was designed to explore these questions.

The achievement and preference findings for the experimental study are perhaps best interpreted in relation to other similar studies. A comparison of the three studies that have had similar treatments, content, classroom contexts and achievement criteria seems to yield a completely contradictory picture, with results favoring low control for Shymansky and Matthews (1974), high control for Crocker, Bartlett and Elliott (1976) and no significant differences for the present study. An additional study, with similar treatments and content but within a recitation mode (Program on Teaching Effectiveness, 1976) showed results favoring high control using a specific achievement criterion in which the teacher was the source of information. Closer examination of these studies, however, suggests the interesting hypothesis that a gradual shift may occur over time, with low control yielding more positive results only after a longer period. Thus, the two short term studies (Crocker, Bartlett and Elliott and Program on Teaching Effectiveness) favored high control, the intermediate term study (the present study) showed no significant differences, while the longest study (Shymansky and Matthews) yielded differences in favor of low control. It would appear that it would be worthwhile to design a study specifically to investigate this hypothesis.

If the two preference criteria are taken together, it seems reasonable to conclude that pupils do show a strong tendency to desire to work in a low control setting. Again, this is contrary to the finding of the Crocker, Bartlett and Elliott study, suggesting at least the possibility that a shift in preference may occur before a shift in achievement. This seems reasonable if one argues that pupil role perceptions influence outcomes and that shift in role perception might be expected to precede a shift in outcomes. Again, the presence of interaction effects both complicates and refines our view of the influence of treatments. In this case, however, unlike achievement, the main effect could be considered to be relatively independent of the interaction effects. That is, the interactions were ordinal rather than disordinal in nature.

The treatment-teacher interaction is particularly interesting since it suggests that teachers may have a greater impact on pupils' feeling towards the treatments than on pupil performance. It also raises the question, of course, of the degree to which the teachers differed in their implementation of treatments, in spite of the attempt to define teacher behaviors and to train teachers in these behaviors. As indicated in Chapter 6, such differences did, in fact, exist. However, there seems to be no direct connection between teacher behavior differences and the specific form of the treatment-teacher interactions. The possibility does exist, of course, that teachers may have been different on behaviors relevant to the treatments but not included in the operational definitions. In particular, it might be hypothesized that teachers may have differed in the degree to which one of the treatments was congruent with their "natural" teaching style. It is interesting to note that, as perceived by the investigators, one of the teachers appeared to be strongly inclined towards high control. However, this did not appear to reflect in a preference for high control by her pupils. On the other hand, Teacher 1, who evoked the most extreme differences in preference on the part of pupils, seemed to be more at home in low control, at least to the extent of being more tolerant than others of the apparent greater degree of disorganization in the classroom associated with the low control treatment. In any case, the analysis suggests that teacher differences are of substantial importance. This has particular implications for methodology since it suggests that studies employing intact classes, using regular class teachers, will inherently confound two potentially important effects, those of class and teacher characteristics.

At this stage it appears premature to attribute major importance to any of the specific ATI effects. Rather, it is likely that some complex composite of pupil characteristics, or some set of high order interactions may exist. Intercorrelations of the pupil characteristic variables influence the order in which such variables are entered in a stepwise regression analysis. The effect of one variable may thus be subsumed by a previously entered variable or, more precisely, both such variables may reflect the same underlying trait. In contrast, the larger number of significant interactions for the choice task may have occurred because the intercorrelations of variables were not accounted for in the analysis. In any case, for achievement, the interactions accounted for substantially more of the variance than did treatments, although in practical terms neither treatment nor interactions contributed a particularly large proportion of variance. For the preference scale, again interactions were, in general, more significant than treatments. Here, however, the interactions accounted for much more of the variance than was the case for achievement (18 and 13 percent for ATI's and 14 and 12 percent for teachers for the two rounds). In general, it seems reasonable to conclude that for achievement the predominant influence is pupil ability while for preference interactions with pupil personality and with teachers are most important.

The findings for classroom behavior variables must be interpreted in the light of the relatively small amount of pupil-pupil discourse observed in the videotaped lessons and the large within-cell variance in the frequencies of the various categories. Although clear treatment, teacher and treatment x teacher interaction effects did emerge, some of these accounted for particularly large proportions of variance in classroom behaviors. At the same time, if one accepts the premise that pupil perceptions of treatments, particularly as expressed in the preference instruments, stem from classroom processes, then the classroom behavior results help explain the existence of differences in preference. It does seem plausible to suggest that pupil preference differences may well result simply from perceived differences in freedom to engage in pupil-pupil talk.

The low frequency of pupil talk (averaging about nine sentences for low control and five sentences for high control in a nine-minute time sample) is striking, particularly since data from the descriptive study suggested that pupil talk would average in the vicinity of 40-50 sentences in the same time. This immediately indicates that the lessons taught in the experimental treatments were not comparable to those taught in the more natural setting of the descriptive study. While there is no clear evidence on this point, there seems to be no strong basis for attributing this difference to differences in teachers, pupils, lesson content, or other contextual variables. It is more likely that the explanation lies in the attempt to explicitly define teacher behaviors in the experimental treatments. Lessons recorded in the descriptive study were taught in the teacher's natural style. Those in the experimental treatment were taught according to an externally imposed structure. Throughout the experiment, it was observed that, even in low control, teachers tended to remain the focus of interaction. It is possible that, in attempting to implement the treatments as faithfully as possible, teachers were reluctant to allow much pupil-pupil interaction. While this may appear to contradict the definition of low control, it is noted that both treatments were defined in terms of teacher rather than pupil behaviors. A teacher could well have interacted with pupils in a manner consistent with low control while at the same time failing to provide opportunity for extensive pupil-pupil interaction. None of the variables used in monitoring the implementation of the treatments was capable of yielding data on this possibility. It might be noted, however, that teachers did make efforts to develop lesson scripts which, in their perceptions, were consistent with the treatments as defined. It seems likely that attempting to follow a script, even one written for low control, would lead to a strong focus on the teacher throughout the lessons.

The results of the pupil interviews suggest one possible explanation for the failure of the treatments to influence the achievement criterion. If pupils, in general, did not perceive the purposes and procedures in the lesson in process terms, then it is possible that they were unable to relate the test items to what had been learned in the lessons. Under these conditions, pretest-posttest gains would have to be attributed largely to maturation. The interaction of treatment and pupil attitude might then be due to some exception to the general case in which pupils at a particular level of attitude may have perceived one treatment more strongly in process terms than the other.

The interview data also tends to reïnforce the conclusion of the descriptive study that teacher and pupil roles are closely differentiated and that the pupil perceives the teacher as the dominant influence in the classroom. This point is particularly notable in view of the fact that pupils' only exposure to the experimental teachers was during the science classes. Either pupils applied a stereotype of teacher and pupil roles to the new teachers or else these teachers, in fact, conformed to the role models already available to the pupils. In either case, the results indicate that more drastic changes than those used in the experimental study will be required before the traditional teacher and pupil roles undergo a change. This study suggests that a change in the direction of less pervasive teacher influence may be easier to advocate than to implement. (Other recent studies have cast doubt also on whether such a change is beneficial).

A Behavior Setting Interpretation

The finding that teacher behaviors remain quite stable across variations in group size within this study and across classroom contexts when this study is compared to other studies, can be explained in terms of the overall concept of a behavior setting. Essentially, a behavior setting is defined in terms of the roles of individuals rather than in terms of surface structural features such as group size or physical location. In this sense, a teacher interacting with a group of pupils constitutes a behavior setting in which teacher and pupil roles have become clearly differentiated and stable as a consequence of substantial experience by both teacher and pupils in this setting. As Dreeben (1970) argues, the authority position of the teacher is an inherent feature of the nature of school organization and the goals of schooling. The teacher knows his role, pupils know their roles, and all behave accordingly. Under these circumstances it is not surprising, for example, that the teacher behaves similarly whether interacting with two pupils or with the whole class (and presumably with any group size in between).

The nature of pupil-pupil interactions, in the absence of the teacher, poses a somewhat different problem. If, as seems plausible, the opportunity to engage in such interactions in the classroom is a relatively unfamiliar experience, then individual roles in this type of setting are less clearly established than is the case in the presence of the teacher. (Pupils are, of course, accustomed to working alone in the classroom, when engaged in seatwork, for example, but this situation is drastically different from the pupil pair setting). Pupil role vis à vis the teacher is evident in the change in pupil behavior when the teacher intervenes and in the tendency for groups to be isolated from each other (since the group structure was defined by the teacher). Beyond this, it seems plausible to suggest that pupils may bring to the group setting a set of role perceptions similar to those that exist if the same pupils were interacting outside the classroom. While there is no direct data on, say, the interactions of children at play, it does not seem especially far-fetched to suggest that such interactions would be found to be characterized by the "ordering around" and the negative ratings that seem to be the most clearly identified aspects of pupil-pupil interaction in this study. In any case, this would form an interesting area of study.

This analysis bears some similarities to that of Cusick (1973) on the relationships between students in high school classes. Cusick observed that high school students fulfilled their expected roles when called upon to interact with the teacher but that discussion among students themselves, within the classroom, simply followed the concerns of the student group in their out of class behavior. In behavior setting terms, the student group, even when it existed in the classroom, formed a different behavior setting from the teacher-class unit. The parallel cannot be carried too far in this case because the groups used in this study were not formed on the basis of pupil affiliations outside the classroom (although the use of proximity of seating as the basis for forming groups may have had this as a consequence), and because the concerns of elementary school pupils are not the same as those of high school students. Unlike Cusick's high school students, the talk of pupils in this study was almost exclusively on-task. Nevertheless, the largely non-substantive nature of pupil-pupil interactions suggests that lesson goals were of less concern to these pupils than were procedural matters.

In considering the results of the experimental study in behavior setting terms, it must be argued that the salient feature of the distinction between the two treatments is the increased opportunity for pupil-pupil interaction in the low control treatment. Even though this opportunity was not as great in the experimental study as in the descriptive study, pupils did interact with each other nearly twice as much in low control as in high control. Pupils may well have preferred low control because of its greater opportunity to operate in a teacher free setting. At the same time, what the pupils do in such a setting does not seem to be very different for the two treatments nor does it seem to matter in achievement terms. In any case, the pupils did not seem to perceive the purposes of the lesson in terms of the goals established by the curriculum being studies, nor did they see any fundamental change in teacher role in the two treatments. This suggests the possibility that the treatments may not have been sufficiently differentiated to yeild the shift in role perceptions that may be required for substantial behavior changes to occur.

Remarks on Methodology

It is always relatively easy, in retrospect, to make suggestions for improvements in the methodology used in a research project. (Critics of a study are, of course, particularly adept at this). It is fairly obvious, for example, in this study that treatments could have been more strongly differentiated, that the use of split classes or similar devices could have brought about some direct control of class effects, that more feedback could have been given to teachers on their implementation of the treatments, that other achievement criteria, including retention and transfer tests could have been used, that more detailed interviews could have revealed more of the nature of pupils' thought processes, and so on. The purpose of this section is not to dwell in detail on such possibilities but rather to raise a number of more general issues about the conduct of this type of research.

First, it became apparent as the classroom observation system was under development that the system could have benefited from a closer analysis of the logical connections between the eventual aims of the study and the observational categories to be used. This matter seemed to be well in hand at the outset, with the substantive-logical categories being used to analyze scientific thinking by pupils as the lesson was in progress, and with various "instructional" sub-dimensions designed to capture the control aspect of the lessons. However, it was soon found that few pupil behaviors could be categorized in substantive-logical terms, although it was not until the interview data were examined did it become apparent that pupils tended not to perceive the lessons in science process terms. Similarly, many of the categories borrowed from the Bellack system simply failed to occur in the classes at hand or, more important, failed to have any clear relationship to the issue of control. Greater forethought, coupled perhaps with some form of cross validation study to determine whether, indeed, the categories were meeting their objectives would have no doubt led to a "leaner" category system with fewer dimensions and categories but in which each category had a justifiable purpose.

In this respect, of course, the study was no worse than most of the recent major classroom observation studies, in which almost every conceivable category was coded and all categories correlated with outcomes. Thus, the point of the above discussion is applicable to studies of that type generally.

The question of the use of videotapes for recording classroom events is one that was of concern to the investigators throughout the study. Although there was little evidence that videotaping seriously disrupted classroom events (references to recording equipment and personnel were coded and were found to be quite rare), the more important question was whether the videotapes were capable of capturing, in a holistic manner, the essence of the lessons. There seems to be no doubt that the time sampling in the experimental study yielded a fragmented picture of both teacher and pupil behavior. On the other hand, an attempt to record, for example, all teacher actions in a single lesson would have presented major problems of time in coding.

In retrospect, it is now felt that it would have been much more efficient to have used an observer in the classroom as an alternative to videotaping. Such an observer could, for example, have coded teacher behaviors on one occasion and the behaviors of selected pupils on other occasions. This would have forced the investigators to reduce the number of categories to a manageable level and would have eliminated the tedious and time-consuming task of developing transcripts and playing and replaying tapes.

A more serious issue is that of class and teacher effects as confounding factors in experimental studies using natural classes. Cost and logistical problems of conducting studies with sufficiently large samples to adequately balance these effects are likely to prohibit such studies. The choice, then, is between "parched up" designs using combinations of sampling and analytical means of overcoming these problems and laboratory studies which, although highly controlled, suffer from problems of generalizability. After having carried out several experiments of the first type (Crocker, Bartlett and Elliott, 1976; Griffiths and Crocker, 1976;

Merricks and Crocker, in press; the present study), the authors are convinced that greater attention must be paid to internal validity. That is, sampling and design considerations must be such that it is possible to attribute specific results unambiguously to treatments. As it happens, this requirement need not be totally incompatible with the generalizability requirement. For example, the assignment of teachers to more than one class, as was done in this study, does not seriously reduce generalizability as long as the teachers used are representative of elementary school teachers. Similarly, reassigning pupils to classes for experimental purposes is not a major threat to generalizability if class size is maintained and if the new groups are together for a reasonable length of time. It is important, of course, in designing more laboratory oriented studies that the integrity of the treatments be maintained. In fact, this must be considered as the crucial element of design. If the findings are to be generalized to natural classrooms, the treatments must be such that they can be used by regular teachers under normal classroom circumstances. This does not entirely preclude "idealized" treatments as a route to understanding basic processes. It does, however, suggest that their applicability is somewhat limited.

In general, it is argued that experimental studies should be more highly focused and should be carried out only if the experimenter has the resources and has obtained sufficient control over the school settings in which he is to work that he is able to implement an internally valid design. Although this would tend to eliminate all but relatively large scale studies, a trend towards such studies is considered to be desirable in any case. Programmatic research rather than isolated small scale studies is essential if the pace of progress in research on teaching is to be accelerated.

Footnotes

[1] The terms "process" and "inquiry" are being used here in a general sense to refer to curriculums characterized by an emphasis on science as a way of acquiring knowledge, the absence of textbooks, the requirement for first hand experiences by pupils in apparatus manipulation, data collection, data interpretation, and so on. The term "process" is given a highly specific meaning in this study, in the context of the classroom observation system and the measurement of outcomes. The term "inquiry" is avoided throughout this report since it has been used to denote both a teaching method and a characteristic of science and has thus confused rather than clarified discourse about science curriculum.

[2] This curriculum (Crocker, 1973) was structurally similar to the AAAS program Science - A Process Approach, but with less emphasis on the hierarchical structure of science processes. Materials consisted of Teaching Guides and Apparatus Kits, as is typical of most of the "new curriculums" in science.

[3] For reviews of these studies, see Flanders (1970) and Dunkin and Biddle (1974).

[4] The intervention sample was confined to ten lessons because data on the remaining three were not available in time for inclusion in the M.Ed. Thesis (Wayne Oakley) based on the analysis of intervention data.

APPENDIX A

INSTRUMENTS

PRE-TEST

PROCESS ACHIEVEMENT TEST

FORM I

INSTRUCTIONS:

1. Do not open this booklet until you are told
2. Please do not write on this booklet
3. Write all your answers on your answer sheet.

e.g. If you think that (a) is the correct answer to Question 1,
Put a in the box on your answer sheet ——— next to Q 1.

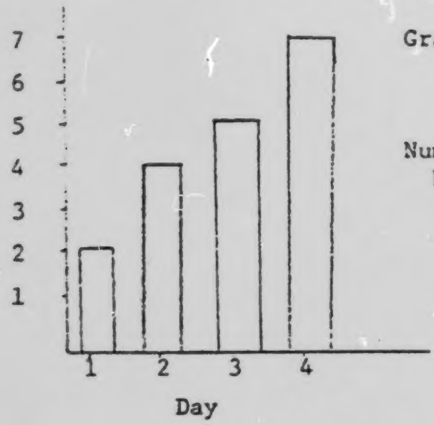
a

1. Homing pigeons are birds that can find their way home whey they are released a long way from home. A man took his two pigeons A & B 20 miles from home and released them. He did the same thing four days in a row. These are the results he got.

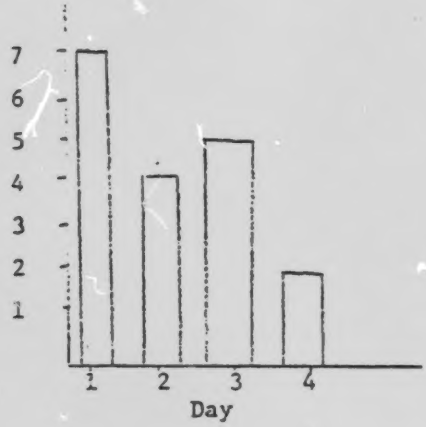
Day	Pigeon A	Pigeon B
1	7 hours	6 hours
2	5 hours	2 hours
3	4 hours	3 hours
4	2 hours	4 hours

Look at these histograms (graphs).

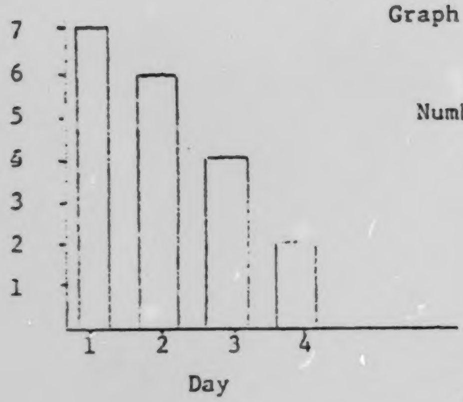
Graph I



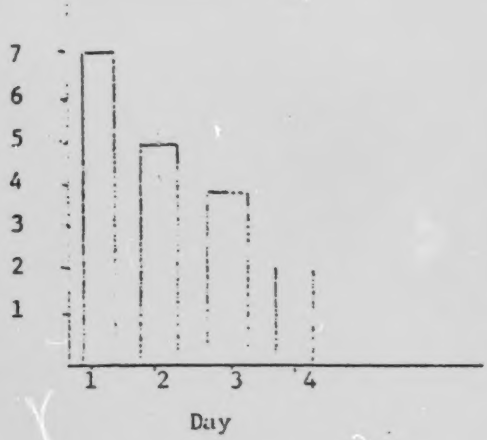
Graph II.



raph III.



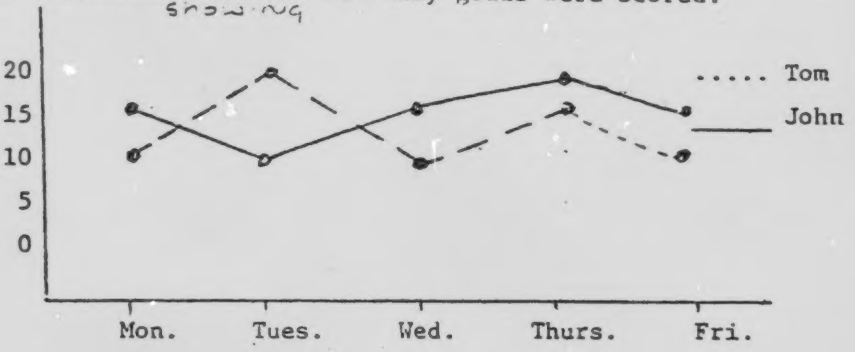
Graph IV.



Which histogram represents the results for pigeon A?

a) Graph I
b) Graph II
c) Graph III
d) Graph IV

2. In hockey practice, Tom and John shot 20 shots each day for five days. The graphs show the number of goals scored each day. Here are the results ~~in terms of~~ showing how many goals were scored.



Look at the graph.

Which of the following statements is correct?

a) John scored more goals than Tom on Monday and Friday
b) Tom scored fewer goals than John on two days
c) John and Tom scored the same number of goals on Tuesday
d) Tom scored more goals than John on Friday

This table shows the highest temperature reached for each day for the first and second weeks of February.

	Week 1	Week 2
Mon.	1	5
Tues.	2	4
Wed.	0 3	3
Thurs.	3	2
Fri.	2	3
Sat.	1	2
Sun.	1	2

3. What is the mean (average) highest temperature for the second week?

   a) 2
   b) 3
   c) 4
   d) 1

4. Which week has the largest range of temperatures?

   a) Week 1
   b) Week 2

5. A group of students wanted to see which color of construction paper becomes warmest in sunlight. Look at the list of variables below and check those you think might influence their results.

   a) [✓ size of the piece of paper]

   b) [✓ the color of the paper]

   c) the type of paper

   d) [✓ where the paper was left]

   e) all of the above

6. A car tire company wanted to know if their new tires would last as long as their old ones. If you were carrying out such a test which variables do you think it would be most important to keep under control?

- a) the time of day the test was made
- b) the number of miles travelled by each tire
- c) the weather conditions
- d) the age of the driver.

7. Two identical cubes of potato A & B were cut from a large potato and weighed. Cube B was then cut into 4 pieces. Cube A and the four pieces of cube B were set aside in the same place. After 3 hours cube A and the four pieces of cube B were weighed again to find how much weight cube A & B had lost.

In this experiment which variables ~~were~~ was changed by the experimenter

a) time between first and second weighings

b) weight of the cut and uncut cubes after 3 hours

c) whether the cube was cut up or not

d) loss of weight of the two cubes A and B.

In Question 7

8. Which variables ~~were~~ was held constant?

a) time between first and second weighings

b) weight of the cut and uncut cubes after 3 hours

c) whether the cube was cut up or not

d) loss of weight of the two cubes A and B

9. In the experiment in Question 7 which variable changed as a result of the experimenter's actions.

   a) time between first and second weighings

   b) weight of the cubes at the beginning

   c) loss of weight of the two cubes A and B

   d) whether the cube was cut up or not

10. Which variables were not controlled but might have influenced the results?

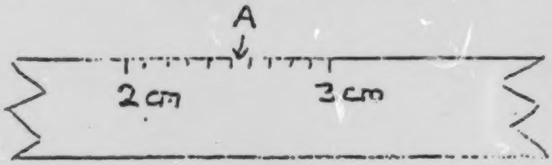
   a) the temperature of the air

   b) air currents

   c) time between weighings

   d) both a and b

11. Look at this portion of a metre ruler. What is the measurement at point A?



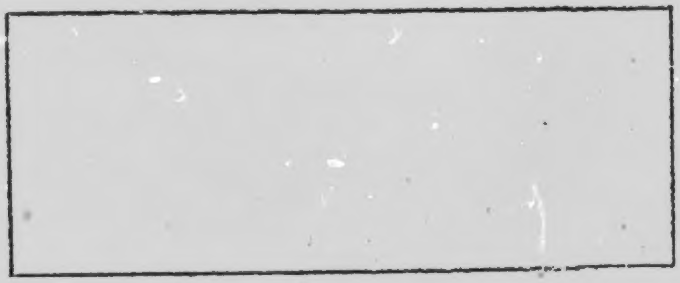
a) 2 centimetres 5 millimetres

b) 2 millimetres 5 centimetres

c) 2 centimetres 6 millimetres

d) 3 centimetres 4 millimetres

12. Using the metric ruler on your desk measure the long side of this rectangle to the nearest centimetre.



The length is:

a) 10 centimetres
b) 9 centimetres
c) 8 centimetres
d) 11 centimetres

13. In the metric system

1 metre = 100 centimetres

1 centimetre = 10 millimetres

Convert

144 centimetres into metres and centimetres

The correct answer is:

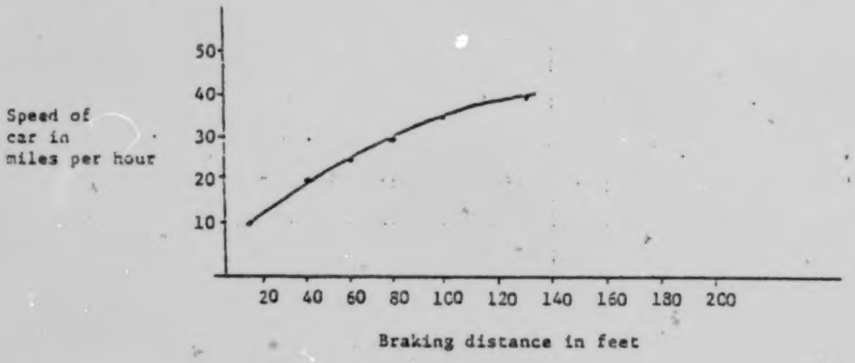
a) 1 metre 14 centimetres

b) 14 metres 4 centimetres

c) 1 metre 44 centimetres

d) 1 metre 44 millimetres

This graph shows how many feet it takes a car to stop, after braking (braking distance), when the car is travelling at different speeds.



14. 1. Predict the braking distance at 15 miles per hour.

a) 20feet
b) 30feet
c) 10foot
d) 40feet

15. 2. Predict the braking distance at 44 miles per hour.

a) 140feet
b) 150feet
c) 200feet
d) 120feet

16. Look at the graph in question
If the car needs 90feet to stop after braking what speed was it travelling?

a) 32 miles per hour
b) 34 miles per hour
c) 31 miles per hour
d) 36 miles per hour

17. Why do you think the prediction for the braking distance for 15 miles per hour is better than that for 44 miles per hour?

a) It is more difficult to measure braking distances at high speeds.
b) The prediction for 15 miles per hour is between two measured values.
c) The prediction for 15 miles per hour is outside any of the measured valu
d) It is easier to measure braking distances at high speeds

POST-TEST I

PROCESS ACHIEVEMENT TEST

FORM I

INSTRUCTIONS:

1. Do not open this booklet until you are told
2. Please do not write on this booklet
3. Write all your answers on your answer sheet.

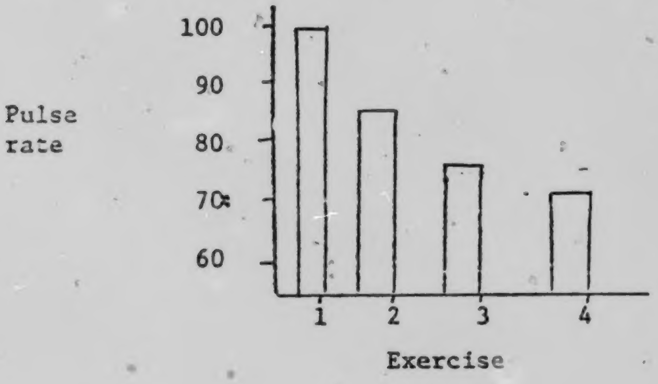
e.g. If you think that (a) is the correct answer to Question 1,
Put a in the box on your answer sheet ——— next to Q 1.
a

1. An experiment was carried out to compare the pulse rates of two persons A and B immediately after they finished three minutes of an exercise. The pulse rates were taken after lying down, typing, walking, running up and down stairs. Here are the results.

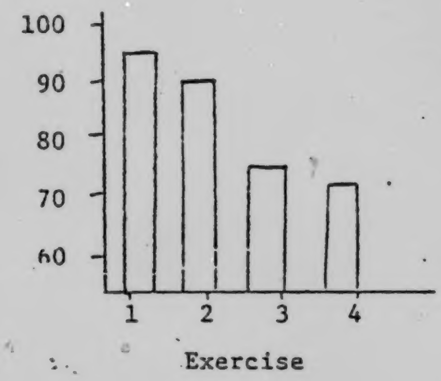
Exercise	Pulse rate of Person A	Pulse rate of Person B
1. Lying down	70	72
2. Typing	75	75
3. Walking	85	90
4. Running up and down stairs	100	95

Look at the histograms below.

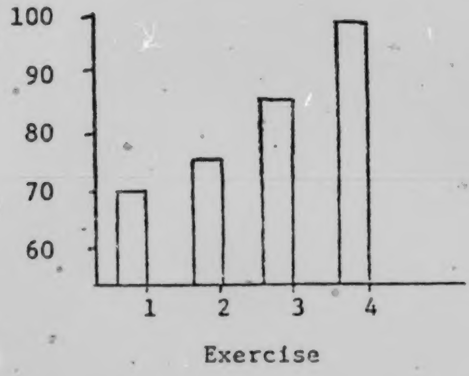
Histogram I



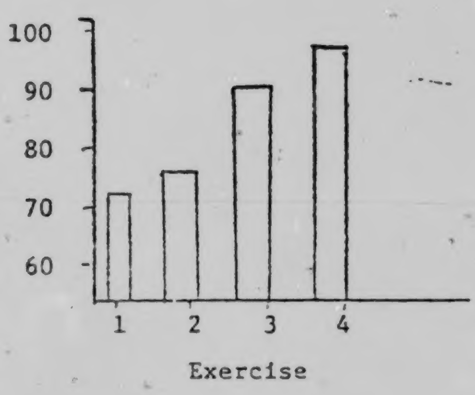
Histogram II



Histogram III



Histogram IV



Which histogram represents the results for person B?

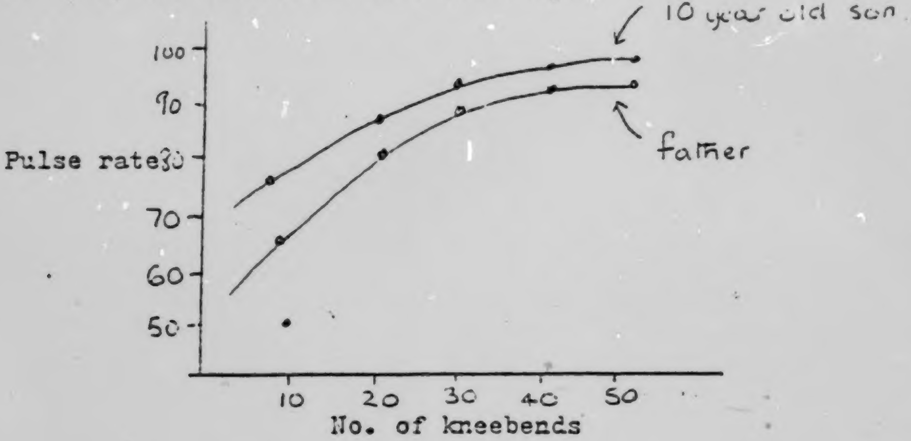
a) Histogram I

b) Histogram II

c) Histogram III

d) Histogram IV

2. The graph below shows the change in pulse rate of a father and his 10 year old son while they were doing 50 kneebends. The pulse rate was measured every ten kneebends automatically.



Look at the graph.

Which of the following statements is correct?

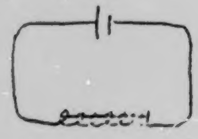
a) The son's pulse rate increased more during the exercise than did his father's.

b) The pulse rate of the father increases more during the exercise than the pulse rate of the son.

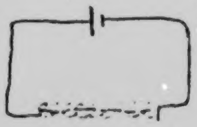
c) The son's pulse rate after 50 kneebends is twice that of his father.

d) Both the father and son have the same pulse rate at the end of the experiment.

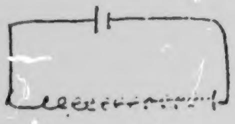
3. An experiment was done to see if an electromagnet could be made stronger by putting more coils in the wire as in th4 diagram below.



6 coils



9 coils



12 coils

Each electric magnet in the diagram above was used to pick up paperclips. In each case several trials were made and these are the results.

No. of coils	No. of paper clips picked up each time					
	Trial 1.	2.	3.	4.	5.	6.
6	1	2	2	3	3	1
9	4	4	5	5	2	4
12	7	6	6	6	4	7

3. What is thw mean (average) number of paper clips picked up with 6 coils?

   a) 6
   b) 5
   c) 2
   d) 4

4. Which number of coils picks up the greatest range of paper clips?

   a) 6
   b) 9
   c) 12
   d) 6 or 9

45 A car tire company wanted to know if their new tires would last as long as their old ones. If you were carrying out such a test which variables do you think it would be most important to keep under control?

a) the time of day the test was made
b) the number of miles travelled by each tire
c) the weather conditions
d) the age of the driver.

6. A scientist wanted to compare the pulse rates of people involved in different jobs, for example, office worker, road digger, doctor, etc. If you were carrying out such an investigation, which variable do you think it would be most important to keep under control?

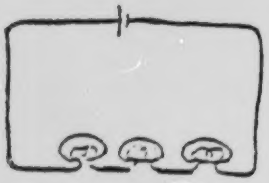
   a) the air temperature

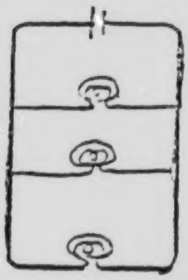
   b) the age of the people

   c) the kind of job that they do

   d) the sex of the people

An experiment was set up to compare the brightness of Three bulbs when they were connected up in series and in parallel as shown in the diagram below.





In an experiment such as this

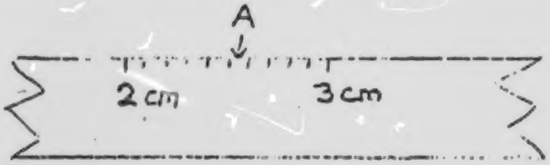
7. which variable was changed by the experimenter?
   - a) number of batteries.
   - b) number of bulbs.
   - c) Type of circuit (series or parallel)
   - d) brightness of the bulbs.

8. In the above experiment which variables were held constant?
   - a) number of batteries.
   - b) number of bulbs.
   - c) Type of circuit (series or parallel)
   - d) both a) and b).

9. In the above experiment which variable would you measure to see if it had changed?
   - a) number of bulbs.
   - b) brightness of the bulbs.
   - c) both a) and b)
   - d) none of the above.

10. Which variables were not under control but which may have altered the results?
    - a) The temperature of the air.
    - b) The distance apart of the bulbs.
    - c) The strength of the battery.
    - d) Both b) and C)

11. Look at this portion of a metre ruler. What is the measurement at point A?



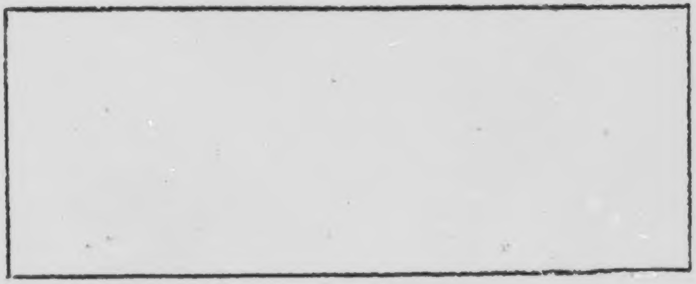
a) 2 centimetres 5 millimetres

b) 2 millimetres 5 centimetres

c) 2 centimetres 6 millimetres

d) 3 centimetres 4 millimetres

12. Using the metric ruler on your desk measure the long side of this rectangle to the nearest centimetre.



The length is:

a) 10 centimetres

b) 9 centimetres

c) 8 centimetres

d) 11 centimetres

13. In the metric system

1 metre = 100 centimetres

1 centimetre = 10 millimetres

Convert

144 centimetres into metres and centimetres

The correct answer is:

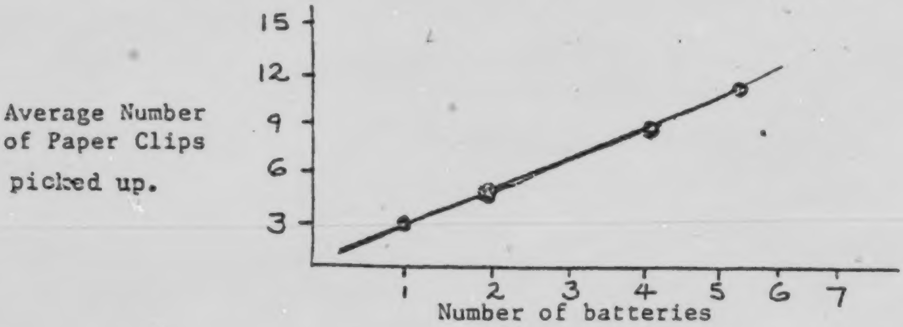
a) 1 metre 14 centimetres

b) 14 metres 4 centimetres

c) 1 metre 44 centimetres

d) 1 metre 44 millimetres

The following graph shows how increasing the number of batteries in a circuit increases the strength of an electromagnet. The strength of the electromagnet is measured by the average number of paper clips that it can pick up.



14. Predict from the graph the average number of paper clips picked up when there are 3 batteries in the circuit.

    a) 5

    b) 7

    c) 3

    d) none of the above.

15. Predict from the graph the average number of paper clips picked up when there are 7 batteries in the circuit.

    a) 12

    b) 15

    c) 10

    d) none of the above

16. Why do you think the prediction for the average number of paper clips picked up when there are 3 batteries is better than the prediction for 7 batteries?

    a) It is more difficult to measure strength of an electromagnet with a large number of batteries.

    b) The prediction for 3 batteries is between two measured values.

    c) The prediction for 7 batteries is outside any measured values.

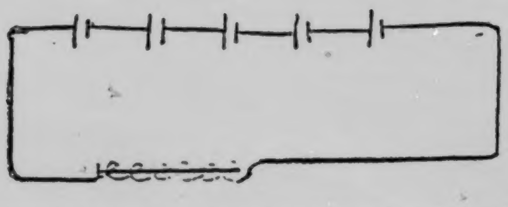
    d) It is easisr to measure the strength of an electromanget with a large number of batteries.

17. Look at the graph on the page before this one. If the electromagnet picks up 7 paper clips how many batteries must there be in the circuit.

a) 4
b) 3
c) 5
d) 2

18. A student did an experiment to see how many paper clips an electromagnet would pick up when there were 5 batteries in the circuit as shown in the diagram below.

He repeated the experiment three times and these are his results.



Results

Trial	No. of paper clips picked up
1.	9
2.	7
3.	5

Since these results vary quite a bit, how could the student best get a more accurate result?

a) take one more reading and see which of the three it is nearest to.

b) take several more readings and average the results.

c) average the three results he already has.

d) none of the above.

19. Your mother sends you to the store to buy enough clothesline to stretch from one end of the garden to the other. The problem is you don't have a ruler to measure the garden. Luckily you find a 'stick' which you can use to measure the garden. You find that the garden is 10 'sticks' long. When you go to the store why must you remember to take the stick with you?

a) so the store can see what your stick is made of
b) so the store can see what shape your stick is
c) so the store knows how long your stick is
d) all of the above

20. If you were to construct a brightness meter, which of the following would be most important?

   a) the length of the brightness meter
   b) the width of the brightness meter
   c) the type of paper
   d) all of the pieces of paper must be the same type.

POST-TEST II

PROCESS ACHIEVEMENT TEST

FORM I

INSTRUCTIONS:

1. Do not open this booklet until you are told

2. Please do not write on this booklet

3. Write all your answers on your answer sheet.

e.g. If you think that (a) is the correct answer to Question 1,
Put a in the box on your answer sheet ______ next to Q 1.
a

1. An experiment was carried out with two groups of students to compare how many words they could remember out of a list of 25 words, when the time between learning the words and being asked to remember them was changed.

Here are the results of the experiments.

Group A

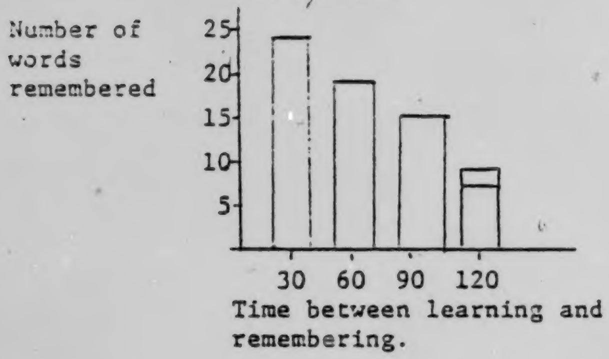
Time between learning and remembering.	No. of words remembered.
30 minutes	23
60 minutes	18
90 minutes	12
120 minutes	7

Group B

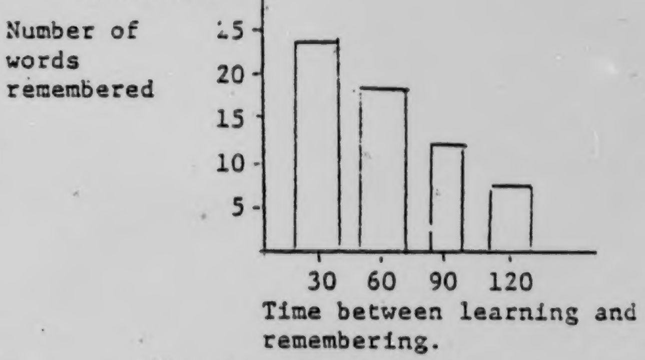
Time between learning and remembering	No. of words remembered
30 minutes	23
60 minutes	19
90 minutes	15
120 minutes	9

Look at these four histograms and answer the question below.

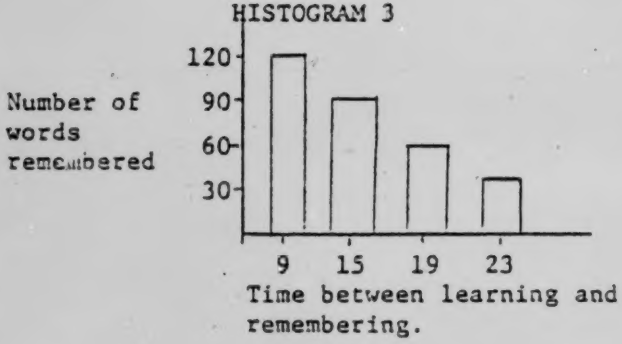
HISTORGRAM I



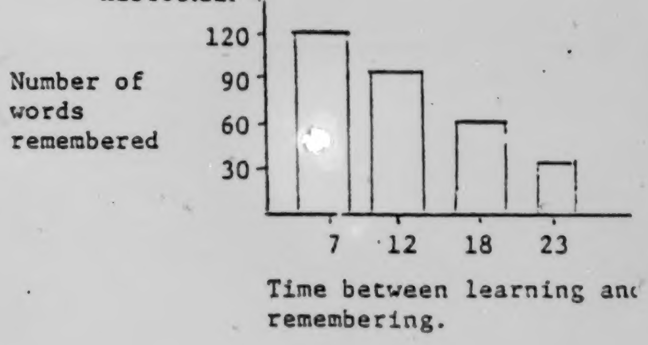
HISTOGRAM 2



HISTOGRAM 3



HISTOGRAM 4



Which histogram represents the data for group A?

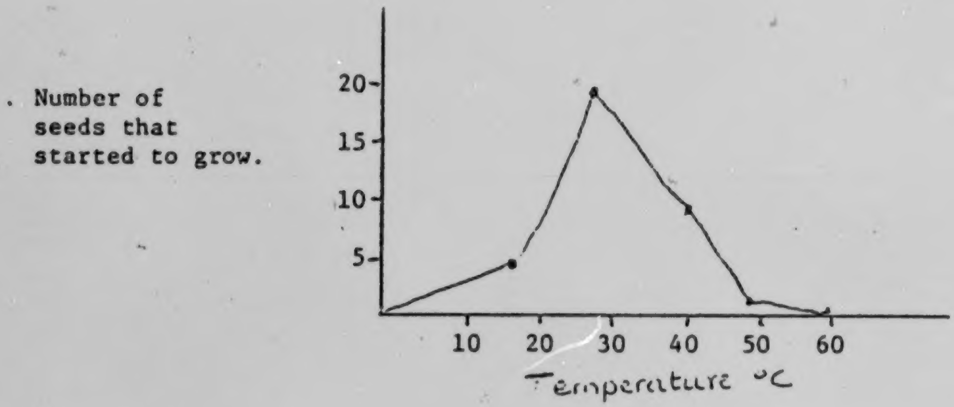
(a) Histogram 1

(b) Histogram 2

(c) Histogram 3

(d) Histogram 4

2. Twenty tomato seeds were planted at each of the following temperatures 10°C, 20°C, 30°C, 40°C, 50°C and 60°C. The following graph shows how many of the 20 seeds started to grow at each temperature.



Look at the graph above. Which of the following statements is correct?

a) The best temperature for tomato seeds to grow is 19°C.

b) The best temperature for tomato seeds to start to grow is 27°C.

c) The best temperature for tomato seeds to start to grow is 15°C.

d) 60°C is the best temperature for seeds to start to grow.

A student was asked to find out whether seedlings growing under a blue light grew taller than seedlings growing under a green light. 10 seeds were planted under a blue light and 10 seeds were planted under green light. The length of each seedling was measured after 10 days. Here are the results:

Seedling #	Under green light	Under blue light
1	6 cm.	5 cm.
2	3	7
3	4	8
4	5	5
5	4	9
6	3	12
7	2	8
8	4	7
9	3	9
10	2	10

3. The mean increase for seedings grown under blue light is:
   - a) 7 cm.
   - b) 8 cm.
   - c) 6 cm.
   - d) 9 cm.

4. Which color light produces the largest range of results?
   - a) blue
   - b) green
   - c) both the same
   - d) impossible to say.

5. A class was carrying out an experiment to find out what variable might alter the time it takes a rod to roll down a sloping table. The class was given four rods A, B, C. D.

   A was 5 cm. long, solid and made of plastic
   B was 15 cm. long, solid and made of metal
   C was 10 cm. long, hollow and made of metal
   D was 10 cm. long, solid and made of metal

   By rolling rods C and D down the same sloping table and timing them, the class could test the effect of which of the following variables?

   a) length of rod

   b) material of rod

   c) solid or hollow

   d) angle of slope

6. In the same experiment as question 5, if you rolled rods A and B down the same sloping table and timed them, the class could test the effect of which of the following variables?

   a) length of rod

   b) material of rod

   c) solid or hollow

   d) angle of slope

A class was interested in finding out whether the age of a person had any effect on the number of words that could be remembered after studying for three minutes.

7. In an experiment such as this which variable would be changed by the experimenter?

   a) the number of words to be learned

   b) the number of words remembered

   c) the age of the person learning the words

   d) the time allowed to learn the words

8. In the above experiment which variable must be kept constant?

   a) the number of words to be learned

   b) the number of words remembered

   c) the age of the person learning the words

   d) both a and c

9. In the above experiment which variable would you measure to see if it had changed as a result of the experiment?

   a) the number of words to be learned

   b) the number of words remembered

   c) the age of the person learning the words

   d) both a and c

10. Which variable was not under control but which may have altered the results?

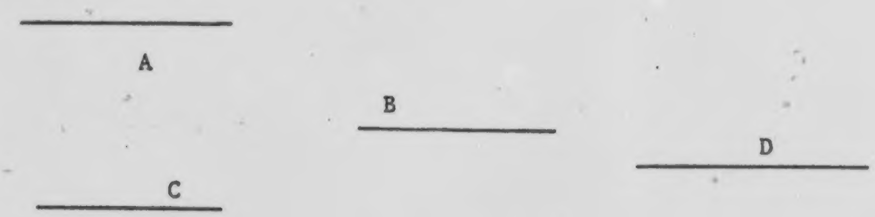
   a) the number of words to be learned

   b) the age of the person

   c) Whether the learners read a lot of books or not.

   d) both b and c.

11. Using your metric ruler measure these four lines to the nearest millimetre.



Which of the following shows the lines A,B,C,D in order of increasing length from shortest to longest?

a) A B C D
b) B C A D
c) C B A D
d) D B C A

12. If you found that your height was 1 metre 50 centimetres, what would be your height in milimetres?

    a) 150 millimetres
    b) 15 millimetres
    c) 1500 millimetres
    d) none of the above.

13. In an experiment to measure the stretch of a rubber band, why would it be more accurate to repeat each measurement several times and average the results?

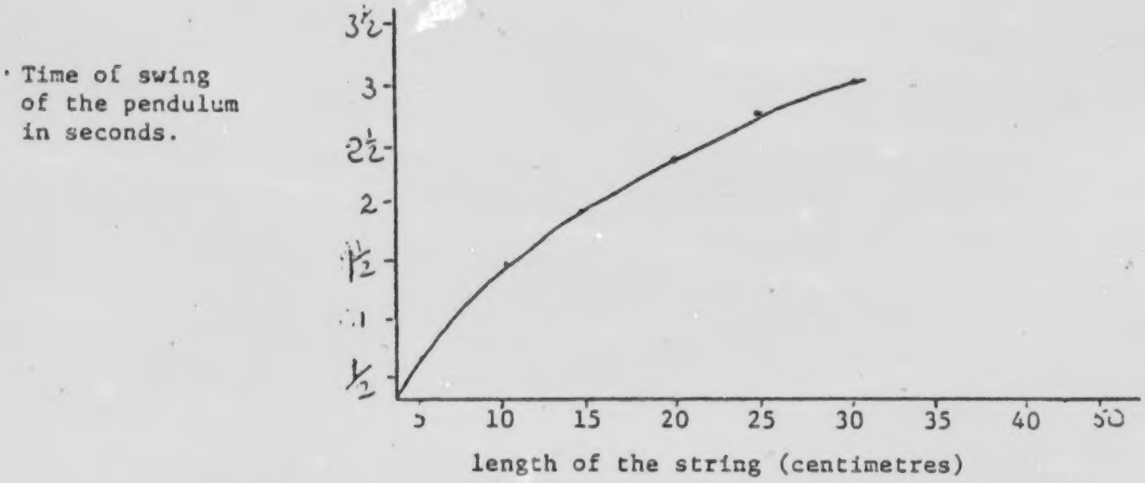
(a) the stretch of the rubber band might vary even with the same weight attached to it.

(b) The longer it takes to get the results the more accurate are the results.

(c) Any measurement errors would be averaged out.

(d) Both a) and c)

A group of students were investigating the effect of changing the length of the string on the time of swing of a pendulum. Here is a graph showing their results.



14 Predict the time of swing of the pendulum when the string is 10 cm.

a) 1½ ~~3~~ seconds
b) 2 ~~3½~~ seconds
c) 1 ~~4~~ seconds
d. 2½ seconds

15. Predict the time of the pendulum when the string is 45 cms.

a) 3½ ~~5~~ seconds
b) 3¼ ~~6~~ seconds
c) 3 ~~7~~ seconds
d) 2½ ~~8~~ seconds

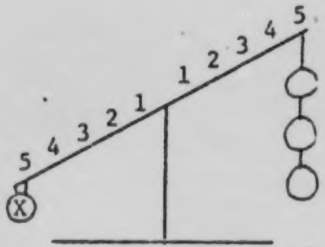
16. What increase in length of the string would change the time of the swing from 3-4 seconds?

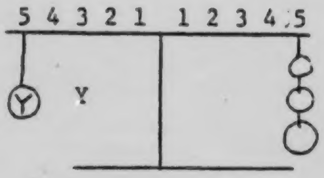
a) 10 ~~seconds~~ centimetres
b) 15 ~~seconds~~ centimetres
c) 5 ~~seconds~~ centimetres
d) 25 ~~seconds~~ centimetres

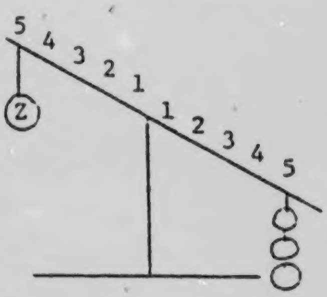
17. Why do you think that the prediction for the time of swing for the 10 cm. long string is better than that for the 45 cm. string?

a) It is more difficult to time the swing with a short string
b) It is more difficult to time the swing with a long string
c) The prediction for 10 centimetres is outside any measured values
d) The prediction for 10 centimetres is within two measured values

18. A student was asked to put 3 objects X, Y and Z in order from heaviest to lightest. Look at the following diagram and answer the question.


5 4 3 2 1
1 2 3 4 5
X


5 4 3 2 1
1 2 3 4 5
Y
Y


5 4 3 2 1
1 2 3 4 5
Z

Which of the following shows the 3 objects in increasing order from lightest to heaviest?

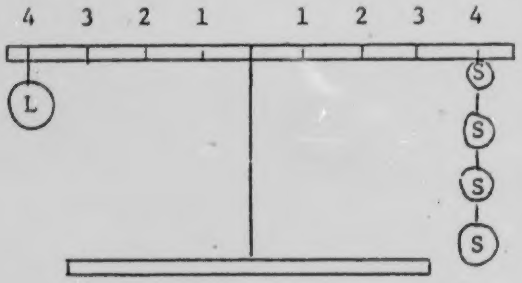
a) Y X Z

b) X Y Z

c) Z Y X

d) Z X Y

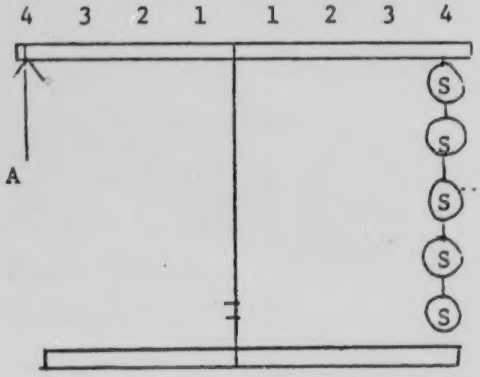
19. Look carefully at the diagram below which shows how many small washers are equal in weight to one large washer.



L = Large washer
S = Small washer

Using the information in the diagram above answer the following question.

If you were given 1 large washer and 4 small washers, which two or three of them would you use at A in the diagram below to balance the washers on the right?



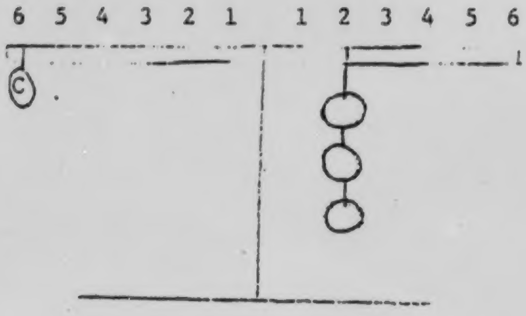
a) 2 large washers

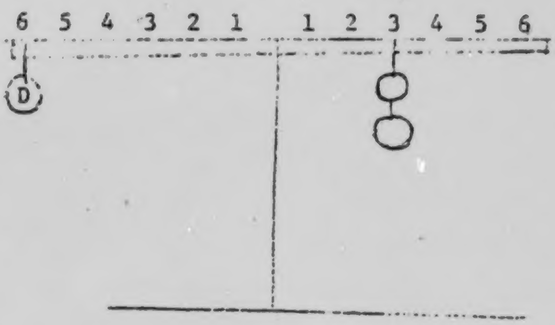
b) 1 large washer and 2 small washers

c) 1 large washer and 3 small washers

d) 1 large washer and 1 small washer

20. Look carefully at the diagrams below in which two objects C and D are balanced with weights.


6 5 4 3 2 1 1 2 3 4 5 6
C


6 5 4 3 2 1 1 2 3 4 5 6
D

Which of the following statements is correct?

a) Object C is heavier than object D.
b) Object D is heavier than object C.
c) Both objects have the same weight.
d) Object D is lighter than object C.

ATSCH & ATSCI

SEMANTIC DIFFERENTIAL

INSTRUCTIONS

Name: ______________________________

On the next two pages are some words which can describe how you feel about YOUR SCHOOL and SCIENCE.

There are four blocks between each pair of words.

Put an X in the block which best describes how you feel.

Here is an example.

Think about how you feel about YOUR SCHOOL.

GOOD ☐ ☐ ☐ ☐ BAD

If you feel your school is really good, mark an X in the block closest to GOOD, like this:

GOOD ☒ ☐ ☐ ☐ BAD

If you feel your school is really bad, mark an X in the block closest to BAD, like this:

GOOD ☐ ☐ ☐ ☒ BAD

If you feel your school is more good than bad, mark an X in the second block from GOOD, like this:

GOOD ☐ ☒ ☐ ☐ BAD

If you feel your school is more bad than good, mark an X in the second block from BAD, like this:

GOOD ☐ ☐ ☒ ☐ BAD

YOUR SCHOOL

Think about how you feel about YOUR SCHOOL.

Mark an X in the block which best describes how you feel about YOUR SCHOOL.

COLUMN A	MOSTLY A	MORE A THAN B	MORE B THAN A	MOSTLY B	COLUMN B
HAPPY	☐	☐	☐	☐	SAD
UNFAIR	☐	☐	☐	☐	FAIR
DULL	☐	☐	☐	☐	EXCITING
USEFUL	☐	☐	☐	☐	USELESS
CONFUSING	☐	☐	☐	☐	CLEAR
EASY	☐	☐	☐	☐	DIFFICULT
PLEASANT	☐	☐	☐	☐	UNPLEASANT
INTERESTING	☐	☐	☐	☐	BORING
AWFUL	☐	☐	☐	☐	NICE
GOOD	☐	☐	☐	☐	BAD

SCIENCE

Think about how you feel about SCIENCE.

Mark an X in the block which best describes how you feel about SCIENCE.

COLUMN A	MOSTLY A	MORE A THAN B	MORE B THAN A	MOSTLY B	COLUMN B
USEFUL	☐	☐	☐	☐	USELESS
FAIR	☐	☐	☐	☐	UNFAIR
EXCITING	☐	☐	☐	☐	DULL
DIFFICULT	☐	☐	☐	☐	EASY
AWFUL	☐	☐	☐	☐	NICE
CONFUSING	☐	☐	☐	☐	CLEAR
GOOD	☐	☐	☐	☐	BAD
UNPLEASANT	☐	☐	☐	☐	PLEASANT
HAPPY	☐	☐	☐	☐	SAD
BORING	☐	☐	☐	☐	INTERESTING

THE WAY I FEEL ABOUT MYSELF

Print your name: First ____________
last ____________

For each question, draw a circle around the letter in front of the best answer for YOU.
There are no right or wrong answers.
Circle just one letter in each question.

1. How well do you think you can do in school compared with your close friends?
   - a. I am the best
   - b. I can do better than most of them
   - c. I am about in the middle
   - d. I can not do as well as most of them
   - e. All of them can do better than I can

2. How well do you think you can do in school compared with the others in your class?
   - a. I am one of the best
   - b. I can do better than most others
   - c. I am about in the middle
   - d. I can not do as well as most others
   - e. Most of them can do better than I can

3. Suppose you were in high school. How well do you think you would do in your class?
   - a. I would be one of the best
   - b. I would do better than most others
   - c. I would be about in the middle
   - d. I would not do as well as most others
   - e. Most of them would do better than I would

4. Forget how others describe your work.
How would you describe your own work?

   a. My work is excellent
   b. My work is good
   c. My work is average
   d. My work is below average
   e. My work is much below average

5. Suppose you had to write a test on all you have learned so far in school.
How well do you think you would do compared with the others in your class?

   a. I would be one of the best
   b. I would do better than most others
   c. I would be about in the middle
   d. I would not do as well as most others
   e. Most of them would do better than I would

6. Suppose you started new school work which was much more difficult than your usual school work. How much trouble do you think you would have with this new work?

   a. I would have no trouble at all
   b. I would have just a little trouble
   c. I would have a fair amount of trouble
   d. I would have a lot of trouble
   e. I would not be able to do the work at all.

PLEASE CHECK TO SEE THAT YOU HAVE ANSWERED ALL THE QUESTIONS

SPI

Print your name: First ______________________ Last ______________________

Read each of the statements listed below very carefully. They are all statements on how you feel about your science lessons. If you agree with the statement underline AGREE. If you disagree with the statement underline DISAGREE. You may ask for help if you do not know a word. Just raise your hand and the teacher will come to your desk.

MARK EVERY ONE.

1. I would like the science lessons better if the teacher gave more directions at the beginning of the lessons. AGREE DISAGREE 1.

2. I would like the science lessons better if the teacher and class talked about the activities less at the end of the lessons. AGREE DISAGREE 2.

3. I would like the science lessons better if the teacher gave her own ideas more, rather than paying attention to what we say about the activities. AGREE DISAGREE 3.

4. I would like the science lessons better if the teacher spent more time asking us to do things, and less time telling us to do things. AGREE DISAGREE 4.

5. I would like the science lessons better if more time was spent on directions (like how to do things, how to behave, and so on) rather than talking about what we worked on in the activities (like batteries and bulbs, variables, data, and so on). AGREE DISAGREE 5.

6. I would like the science lessons better if the teacher spent more time talking to groups, and less time talking to the whole class. AGREE DISAGREE 6.

7. I would like the science lessons better if the teacher came around to groups more often to tell us if we were doing things right or wrong. AGREE DISAGREE 7.

8. I would like the science lessons better if the teacher asked more often for groups to give out their results in class. AGREE DISAGREE 8.

9. I would like the science lessons better if we were reminded more often of the way the teacher wanted us to do the activities. AGREE DISAGREE 9.

10. I would like the science lessons better if we were not reminded so often of what the right answer was in the activities. AGREE DISAGREE 10.

11. I would like the science lessons better if the teacher gave me less time to answer questions she asked me, before she asked someone else or gave the answer herself. AGREE DISAGREE 11.

12. I would like the science lessons better if the teacher gave fewer directions at the beginning of the lessons. AGREE DISAGREE 12.

13. I would like the science lessons better if the teacher and class talked about the activities more at the end of the lessons. AGREE DISAGREE 13.

14. I would like the science lesson better if the teacher paid more attention to what we had to say about the activities rather than giving her own ideas. AGREE DISAGREE 14.

15. I would like the science lessons better if the teacher spent more time telling us to do things, and less time asking us to do things. AGREE DISAGREE 15.

16. I would like the science lessons better if more time was spent talking about what we worked on in the activities (like batteries and bulbs, variables, data, and so on) rather than giving directions (like how to do things, how to behave, and so on). AGREE DISAGREE 16.

17. I would like the science lessons better if the teacher spent more time talking to the whole class, and less time talking to groups. AGREE DISAGREE 17.

18. I would like the science lessons better if the teacher came around to groups less often to tell us if we were doing things right or wrong. AGREE DISAGREE 18.

19. I would like the science lessons better if the teacher asked less often for groups to give out their results in class. AGREE DISAGREE 19.

20. I would like the science lessons better if we were not reminded so often of the way the teacher wanted us to do the activities. AGREE DISAGREE 20.

21. I would like the science lessons better if we were reminded more often of what the right answer was in the activities. AGREE DISAGREE 21.

22. I would like the science lessons better if the teacher gave me more time to answer questions she asked me, before she asked someone else or gave the answers herself. AGREE DISAGREE 22.

DID YOU UNDERLINE ONE CHOICE FOR EVERY STATEMENT? CHECK BACK AND SEE.

Choice Task: Directions for Teachers

In this procedure we are trying to measure student preference for the high or low control treatments.
This measure is to be incorporated into the regular classroom activity.

1. Timing:

At the beginning of the class immediately following activity 1 of the Mechanics Unit.
At this time the students SHOULD KNOW:

- a. how to set up a pendulum
- b. that length, weight, and angle are some variables that could affect the period of a pendulum.
- c. how to time the swings of a pendulum, meaning of "period" of pendulum

****Teachers should avoid telling students how to control these variables in experimenting with the relationships between length, weight and period of the pendulum.

2. Materials:

per student pair: popsicle stick, string, paper clips, washers, tape, access to timer with a second hand.

per student: An index card on which to write his name and indicate his choice.

Instruction sheets: there are two kinds(see below). Each pair of students gets one sheet depending on its choice. Take 20 of each to be sure you have enough.

3. Procedure:

- a. ensure that the students are in their normal groups.
- b. pass out index cards, and have students mark on it their FULL names.
- c. introduction: "In today's lesson we will experiment to see HOW the period of a pendulum depends on the length of the pendulum string and the weight of the pendulum.

  Each one of you can choose between two ways to do this. Think of the one You would LIKE or ENJOY doing. Think for yourself, since you are the only one who really knows what you like doing."

d. explain choices:

Teacher explains: "In the first choice you get the equipment you need, AND a sheet of step-by-step directions you must follow to do the experiment."

Teacher writes on board: 1. teacher gives step by step directions to follow

Teacher explains: "In the second choice you get the equipment you need, BUT you follow directions which you must think of on your own to do the experiment."

Teacher writes on board: 2. you follow directions you think of on your own

e. ask the students to think for a moment, and SECRETLY mark a 1 or 2 on their index card to indicate their choice. Have them choose only one.

f. collect the cards, and put them in a labelled envelope.

g. Teacher: "Perhaps your usual partner chose the one you did not. This is quite all right. We will just switch partners for this activity only."

h. assignment to groups:
  1. ask all those who chose "1" to raise hands.
  2. if both usual partners chose "1" ask them to lower hands.
  3. if any hands remain up, pair off these students. make a group of 3 if number is odd.
  4. pair off the "2's" left by themselves.
  5. Ask all the "1" groups to move to one side,(area, etc.) and "2" groups to go to the other.

i. make sure each group has the right sheet, depending on its choice.

j. make sure each group has equipment.

k. treat students in "1" groups high control; those in "2" groups, low control

l. let them work for 15-20 minutes(this may take you to end of period)

m. take over and proceed as you would normally..incorporating what the students did on this activity into your discussion,etc.

n. Give index cards to ~~Rosetta~~ the investigator at earliest convenience.

CHOICE 1

Follow these instructions carefully.

You are going to look at how the strength of an electromagnet depends on

a. the number of batteries in series
AND b. the number of turns of wire

a. Wind 20 turns of wire on the nail. Keep this the same while you are finding out how the number of batteries changes the strength of the electromagnet.

Make the measurements to fill in this table:

number of batteries in series	strength of magnet (number of clips picked up)
1	
2	
3	
4	

b. Use 2 batteries for your magnet in this part. Keep this the same while you are finding out how the number of turns changes the strength of the electromagnet.

Make the measurements to fill in this table:

number of turns on nail	strength of magnet (number of clips picked up)
10	
20	
30	
40	

Write down how the strength of an electromagnet depends on the number of batteries: ______________________________.

Write down how the strength of an electromagnet depends on the number of turns: ______________________________.

Choice Task for Round 1

CHOICE 2

You are going to look at how the strength of an electromagnet depends on

a. the number of batteries in series
AND b. the number of turns of wire

Make up your own directions and follow them.
Keep your measurements on this paper.

QUESTIONS:
1. How does the strength of an electromagnet depend on the number of batteries?
2. How does the strength of an electromagnet depend on the number of turns?

Choice Task for Round 2

CHOICE 1

Follow these instructions carefully.

You are going to look at how the period of a pendulum depends on

a. the length of the pendulum string

AND b. the weight of the pendulum

a. Put 1 washer on your pendulum. Keep 1 washer on the pendulum while you are finding out how the period of the pendulum changes with the length of the pendulum string.

Make the measurements to fit in this table:

length of pendulum string (centimetres)	Period of pendulum (time for 5 swings)
10	
20	
30	
40	
50	

b. Use a length of 50 centimetres for this part. Keep this the same while you are finding out how the period of the pendulum is related to the weight of the pendulum.

Make the measurements to fill in this table:

Weight of pendulum (number of washers)	Period of pendulum (time for 5 swings)
1	
2	
3	
4	
5	

Write down how the period of a pendulum depends on the length of the pendulum string: ______________________

____

Write down how the period of a pendulum depends on the weight of the pendulum: ______________________

____

Choice Task for Round 2

CHOICE 2

You are going to look at how the period of a pendulum depends on

a. the length of the pendulum string
b. the weight of the pendulum

Make up your own directions and follow them.
Keep your measurements on this paper.

QUESTIONS:

1. How does the period of a pendulum depend on the length of the pendulum string?

2. How does the period of a pendulum depend on the weight of the pendulum?

Pupil Interview Form

Name ______________________

Teacher ______________________

Unit ______________________

Interview No.______________________

School and Class ______________________

Interviewer ______________________

Activity ______________________

Treatment ______________________

1. Did you enjoy doing today's experiment? ☐
   (1) enthusiastic; (2) positive; (3) indifferent; (4) negative;
   (5) highly negative. ______________________________________

2. If you went outside after class and someone asked you what you did in science class today, what would you say? ☐
   (1) procedure description; (2) learning-content;
   (3) learning-process; (4) fun; (5) nothing; (6) negative
   ________________________.

3. What do you think was the purpose or reason for doing this activity? ☐
   (1) procedure description; (2) content; (3) process; (4) don't know
   ______________________________.

4. What do you think you learned from the activity? ☐
   (1) procedure; (2) content; (3) process; (4) don't know;
   (5) nothing; ________________________________

5. Did you get enough information from the teacher in order to get started on today's experiment? ☐
   (1) yes; (2) no; (3) don't know; ________________________.

6. Suppose you had been given "fewer" directions. Would you have been able to do the experiment as well? ☐
   (1) just as well; (2) almost as well; (3) fairly well;
   (4) poorly; (5) not at all; (6) don't know; _____________.

7. Did you have the time to do all the things you would like to have done with the materials in the activity?
(1) yes; (2) no; ____________________ ☐

8. Why do you say that (pertaining to ques. 7)?
(1) did all we wanted to do; (2) did everything the teacher wanted; (3) play; (4) trial and error; (5) content ; (6) process; (7) procedure ________ ☐

9. If you had a choice, would you have done the experiment the same way your teacher asked you to do it, or would you like to have done the experiment another way?

(1) same way; (2) another way; (3) not sure ________ ☐

10. Could you explain your answer (pertaining to ques. 9)?
(1) liked the way the teacher did it; (2) couldn't see any other way; (3) would have followed another procedure; (4) didn't like the way the teacher wanted us to do it;
__________________ ☐

11. Who do you think played the most important part in doing the experiment today, students or the teacher?
(1) teacher; (2) students; (3) both did the same;
________________ ☐

12. Why do you say that (pertaining to ques. 11)?
(1) teacher told us just what to do; (2) students didn't do much; (3) students did all the work; (4) students did experiment; (5) teacher didn't do much; (6) students and teacher did equally important parts; ____________ ☐

13. Who do you think did the better job on the activity, you or your partner?
(1) pupil being interviewed; (2) partner; (3) both did the same; (4) don't know. __________________ ☐

14. Why do you say that (pertaining to 13)? ☐

____

15. Think about how your partner might feel about the activity. ☐
Who does your your partner think did the better job?
(1) pupil being interviewed; (2) partner of student being interviewed; (4) both did the same; (5) don't know

____

16. Why do you say that (pertaining to ques. 15)? ☐

____

17. Were you and your partner (name) able to work well together? ☐
(1) yes; (2) no; (3) indifferent; ______________

18. Could you explain your answer a little more (pertaining to ques. 17)? ☐
(1) we shared work; (2) compensated each other; (3) partner bossy; (4) partner took all the stuff; (5) partner mad because I used to do it all; (6) we are good friends; (7) don't like each other;

____

19. Do you think you did the activity the right way: ☐
(1) yes; (2) no; (3) don't know. __________

20. Why do you say that: ☐
(1) we got the right result; (2) teacher told us
(3) followed direction (4) we understood; __________

21. Do you think you did the activity the way the teacher wanted it done? ☐
(1) yes; (2) no; (3) don't know; _______

22. Do you think it was very important that you do the activity the way the teacher wanted?
(1) yes; (2) no; (3) don't know ______________

23. Do you think you got the right results?
(a) yes; (2) no; (3) don't know; ________________

24. How do you know the results were (right or wrong)?
(1) same as others; (2) teacher said; (3) good experimenting; (4) followed correct procedure; (5) don't know; (6) followed wrong procedure; (7) different from other; (8) apparatus problems. ______________________

25. What do you think these results show?
(1) procedure; (2) process; (3) content; (4) don't know.
____

Omit ✓ 26. Why did you (carry out particular tasks during activities)?
(1) process; (2) content; (3) procedure; (4) teacher said; (5) saw others; (6) partner; (7) _____

Omit ✓ 27. Process Task

I. Measurement	1.S	2.U
II. Controlling Variables	1.S	2.U
III. Interpreting Data	1.S	2.U

APPENDIX B

CODING MANUAL, CLASSROOM OBSERVATION SYSTEM

A CATEGORY SYSTEM FOR THE ANALYSIS OF INTERACTION PATTERNS IN ELEMENTARY SCIENCE CLASSES

Working Paper

(Revised July, 1975)

R.K. Crocker, R.P. Amaria, D. Genge,
W. Oakley, G. Clark

Memorial University of Newfoundland
St. John's, Nfld., Canada

The research reported herein is supported by a grant from the Humanities and Social Science Division of the Canada Council. The category system described is in the developmental stages and should not be used without the permission of the authors.

The category system described is being designed primarily for the purpose of analyzing the wide variety of interaction patterns found in elementary science activities. The system permits the recording of various dimensions of teacher-class, teacher-small group, and pupil-pupil interactions in a classroom setting in which pupils are engaged in the type of small group activities that characterize many elementary science programs. With appropriate modifications, the system should be applicable to other classroom settings in which small group work is the dominant feature.

The overall purpose of the study, the theoretical underlying development and use of the system, and the relation of the system to other interaction analysis systems have been discussed elsewhere (Crocker, 1975). This paper describes in detail the system itself, and outlines a set of coding instructions. Procedures to ensure inter rater reliabilities in coding are also discussed. This document can therefore be regarded as a preliminary coding manual and user's guide. Since the Bellack system (Bellack, 1966) was used as the starting point of this study, potential users should first become familiar with that system.

RECORDING AND TRANSCRIBING PROCEDURES

The category system has been developed through the analysis of videotape recordings of twelve science lessons in grade two, four and six classes. In all classes, the activities were representative of those found in a process based elementary science curriculm (Crocker, 1973) currently in use in Newfoundland schools. The characteristic lesson pattern was for the teacher to introduce the class to some science problem and for the pupils to investigate this problem by working in pairs, each pair having its own set of apparatus. Throughout the class session, the teacher circulated among the pupils, discussing the activity with various groups. A teacher-class summary took place at the end of the activity.

In each class, two pairs of pupils were randomly selected and their activities recorded using two independently operated portable videotape recorders. Teacher-class interactions, which tended to occur mainly during the introductory and summary phases of the lesson, were recorded by both systems. Pupils were aware that recording was taking place, although the specific pupils recorded were not aware until the end of the session that they had been singled out. Following the class session, the four pupils recorded were given tests of academic ability, personality, creativity, developmental level, and reading ability.

Written transcripts were prepared from the videotape using the following procedures:

1. An audiotape was prepared from the original videotape.
2. A preliminary typed transcript was prepared from the audiotape.
3. The videotape was viewed by an editor (who had been present in the classroom during the recording session) and corrections made to the transcript.
4. The edited transcript was retyped, with each sentence beginning on a new line.
5. The videotape was agian viewed and timing marks placed on the transcript at six second intervals. The timing marks were designed to give codes a sense of the pace of discourse, to permit the coding of physical action, and to assist in locating particular passages on the tape. Further editorial changes in the transcript were made at the time.
6. The videotape was viewed a third time and physical actions coded directly onto the transcript.

CATEGORY DESCRIPTIONS AND CODING INSTRUCTIONS

Coding Unit

1. The basic coding unit is the sentence. Judgment on what constitutes a sentence are made by the transcriber and editor as the transcripts are prepared. Final corrections in sentence structure are made by the coders.
2. Partial sentences and sentence that the speaker fails to complete are coded as complete sentences. If the intent of the incomplete utterance is not clear, NOC is coded in the move dimension and the remaining dimensions are not coded.
3. Sentences whose meaning can be made clear from the tape are left blank in the initial coding and are later referred to the tape.
4. Physical actions are ordinarily coded simultaneously with the utterance of

a sentence. If an interval longer than six seconds passes with no verbal utterance, physical actions are coded using the six second interval. Such codes are simply entered on the coding sheet sequentially as physical action only. All other dimensions are left blank. For purposes of analysis, the six second interval is taken as equivalent to twice the time taken to utter a sentence.

5. Physical actions are coded for three persons (the teacher and the two pupils being recorded). If a person cannot be seen on the screen, his physical actions are not coded. The exception to this is if the view of a person is momentarily obstructed and it can reasonably be inferred that his physical action during that time is the same as that occurring immediately before and after.

Lesson Phase

At the outset, it was anticipated that each lesson would consist of a number of distinct lesson phases both from a substantive and an instructional standpoint. On the basis of previous experience with lessons of this type, the following pattern was expected:

1. Teacher introduction of problem (teacher-class interactions predominant)
2. Data gathering (pupil-pupil interactions predominant but teacher-group, teacher-pupil, and teacher-class interactions can occur)
3. Data interpretation (interaction pattern varies from lesson to lesson: certain problems lead to class composite results and teacher-class interactions, others require interpretations within groups and hence lead to more pupil-pupil interaction)
4. Summary (again, teacher-class interactions tend to predominate)
5. Cleanup

It was decided to attempt to associate each unit of discourse with one or other of these lesson phases in order to shed some light on the time taken for each phase (and the variations in time from lesson to lesson) and on the specific nature of the interaction patterns found in each of the phases. The following categories have been used for this purpose:

Problem specification (PRB)
Data Gathering (DAG)

Data Interpretation (DAI)
Summary (SUM)
Cleanup (CLE)

The meaning of the above categories is straightforward. Coding problems occur only in transitions since, by definition - the lesson phases tend to occur over reasonably long periods. The coding instructions give a number of conventions used in dealing with transitions.

Within each lesson phase, specific behaviors are considered to be either directly on task, indirectly on task, or off task. The code numbers 2, 1, 0 are assigned to these degrees of attention to task.

Coding Instructions For Lesson Phase

1. PRB is coded from the beginning of the lesson until it becomes clear that data collection relative to the problem at hand is occurring. Elements of data collected simply to elucidate the nature of the problem become part of the PRB phase. Typically PRB will include the entire teacher introduction plus the early part of the group work.

2. Generally DAC follows PRB as a major lesson phase. However, if pupils engage in restructuring the problem after it has been defined by the teacher or if the pupils attempt to make a broad problem more specific, PRB is used.

3. DAI as a task area does not necessarily coincide with DAI or INT in the S/L dimension. In general, DAI is coded as task area when pupils move from data collection to a phase in which they attempt to draw conclusions. Even if additional data is collected at this time, the phase remains DIA and the appropriate S/L code is used to identify the specific nature of the data collection.

4. DAI is coded for teacher-class lesson summaries when specific reference is being made to the data collected by the pupils. For more general summaries, SUM is used.

Examples:	"It seems that some people got the metal to float while others found that it sank."	DAI
	"We see that metals sometimes float and sometimes sink."	SUM

5. Indirect attention to task is coded only when pupils are engaged in substantive activities other than those assigned by the teacher or when the teacher digresses into a broader substantive area then that identified at the beginning of the lesson.

Example:

What are some other ways in which clocks can be made to keep time? (follow an investigation of a pendulum). SUM/1/....

Speaker

Both the speaker and the person to whom the utterance is addressed are coded using the following numerical codes:

0 Unidentifiable or indefinite (usually receiver but sometimes speaker)

1 Teacher

2, 3 } Pupils in first group under analysis

4, 5 } Pupils in second group under analysis

6 Identifiable pupil extraneous to group under analysis

7 Group member unclear (used when a member of the pair under analysis speaks or is addressed but the specific individual cannot be identified)

Modifications of this numbering system can, of course, be made to accomodate different grouping arrangements.

Coding Instructions For Speaker

1. The application of these code numbers generally is straightforward. For example, 1, 0 is coded for teacher-class utterances 2, 3 or 3, 2 for pupil-pupil utterances for members of the first group, and so on.
2. In the case of two persons speaking simultaneously, only the utterance that seems to contribute most to the ability to analyze the lesson is coded. For example, if the teacher gives a direction to the whole class at the same time that the pupils under study are talking within their group, the teacher would normally be coded as the speaker. On the other hand, if the teacher is speaking and one of the pupils in the group interrupts or otherwise comments, then the pupil utterance would be coded.
3. 6 is coded whenever the speaker or receiver is clearly a single individual, whether or not he can be identified by name.
4. 7, 6 is used when an unidentified group member speaks to an extraneous pupil. 6, 7 is coded when an extraneous pupil speaks to the group.

Pedagogical Move

In the Bellack system, the pedagogical move was defined as the basic unit for both coding and analysis. In developing the present system, it was decided that the sentence was a more appropriate coding unit, partly because it reduced the ambiguity in deciding the beginning and end of a move and, more importantly because it was found that a number of possible categories in the other dimensions of the system could be coded for a single move. In spite of changing the coding unit , however, the concept of a move was regarded as too valuable to discard, particularly since the frequency of occurrence of particular moves by teacher and pupils in conventional classes had been well documented and since it was of some interest to compare the classes under study with more conventional classes.

The four pedagogical moves identified by Bellack are; structuring (STR), soliciting (SOL), responding (RES), and reacting (REA). The meanings of these terms given below follow Bellack (1966), with only minor modifications. Only the specific examples are drawn from the lessons used in the present study. It should be pointed out that direct comparability in moves coded using individual sentences and moves coded in the Bellack sense is possible only by considering all contiguous sentences uttered by the same speaker and coded as the same move to be additive to form a single move as defined by Bellack.

STRUCTURING (STR)

Structuring moves are those that set the overall pattern for a lesson or lesson segment. STR moves serve to change the direction of the discourse or action by (1) launching or halting-excluding interactions or (2) indicating the nature of the interaction in terms of time, topic, regulations, etc. Structuring moves do not elicit responses nor are they in themselves responses.

Examples

1. Now, what we are going to do is take each of the objects in the tray and put them in the water. 1,0/STR

2. Now, when we finish, we will place the objects in two groups, those that sink and those that float. 1,0/STR

SOLICITING (SOL)

Moves in this category are designed to elicit an active response on the part of the receiver, either in the form of a verbal, cognitive, physical, or emotional action. Soliciting moves may take any grammatical form (interrogative, imperative, or declarative). In the case of teacher solicitations the imperative form is more frequent while for pupil solicitations of the teacher, the interrogative form is more common. Soliciting moves are coded in terms of the response expected rather than the solicitation itself.

Example

1.	Teacher, come look at this.	2,1/SOL
2.	What do we mean, Lisa, when we say something floats?	1,6/SOL
3.	Put the big washers in the cup.	1,7/SOL

RESPONDING (RES)

Responding moves occur in reciprocal relationship to soliciting moves. Responses can be elicited only by solicitations. (Statements in response to a STR are coded REA).

Examples

1.	Now what do we mean, Lisa, when we say something floats?	1,6/SOL
	It stays on top of the water.	6,1/RES
2.	Teacher, come look at this	2,1/SOL
	In a minute.	1,2/RES

REACTING (REA)

Reacting moves are occasioned by a SOL, STR, RES, or previous REA but are not directly elicited by these moves. The REA move serves to modify be clarifying, rating, synthesizing, etc. In distinguishing between RES and REA, the intent of the previous move must be determined.

Examples

1.	Try making the string shorter.	1,2/SOL
	OK	2,1/RES
	There, that looks better.	1,2/REA
2.	First we must shorten the string and then count the number of washers	1,0/STR

	I don't think that will work.	6,1/REA
3.	Why don't you try shortening the string?	1,2/SOL
	That's a good question.	2,1/REA

NOT CODABLE (NOC)

A sentence is not codable if it is incomplete to the point that its substance is lost, if more than one influence about its intent is possible from the context, or if more than one more can apply to the same sentence.

Coding Instructions for Pedagogical Move

1. STR and SOL are coded in preference to RES and REA when in doubt since the first two serve to advanance the discourse.
2. Delayed responses to a SOL are coded RES but it must be clear that the RES relates to some previous SOL, otherwise code REA.
3. RES is coded if more than one pupil responds to a SOL, even if one specific pupil has been requested to respond.
4. Parenthetical statements occurring in the midst of a series of STR or SOL moves are coded REA.

Example: Teacher is demonstrating a procedure

Now we see if the rock will balance.	1,0/STR
There you go.	1,0/REA
It takes twelve washers.	1,0/STR

5. Incorrect responses to a SOL are coded RES.
6. A speaker cannot respond to his own SOL. Responses to a rhetorical question are coded REA.
7. Responses to a SOL that take the form of questions (e.g. tentative responses in a questioning tone of voice) are coded RES. However, statements following a SOL which are not directly related to the SOL are coded according to context, and are not necessarily coded RES.

Example:

How many washers does it take?	1,0/SOL
Is it five?	6,1/RES

8. A REA is in progress when the speaker:

   a. Evaluates or otherwise discusses a previous move

b. Rephrases or makes reference to a previous move.

c. Expands or interprets a previous move, or draws conclusions from the same point.

d. Summarizes or reviews a previous move

A REA ends when any of the following occurs:

a. A SOL begins.

b. The utterance ends.

c. A STR begins, such as by some verbal convention such as " Now let's go on to......"

d. A distinct shift occurs to another substantive area.

9. A special case of REA occurs when the particular utterance is elicited by physical, action, by observed phenomena, the act of measurement, etc. If the REA is to something that is observed, PHE is coded under nature of reference. Occasionally the REA is in the form of a teacher rating of some pupil action that has not been recorded.

Example:	It floats.	4,5/REA/.../PHE/
	Look, it floats. (While this example seems similar to the first, it is a SOL because of the calling attention to the phenomenon.	4,5/SOL/.../PHE/
	You shouldn't spill the water.	1,4/REA/.../NEG/

10. Following from the previous rule, however, if the teacher gives a command or request based on some pupil action not previously coded, the move can be coded SOL

Examples:	Don't spill it now	SOL
	Don't start yet	SOL

11. Some ambiguity can arise in the coding of SOL and STR when the teacher is giving directions. In general, if the pupils move about, start manipulating apparatus, etc. following the direction, the direction is coded SOL. If directions are more general and seem to apply to procedures to be carried out later in the session or throughout the session, STR is coded.

Examples:	What we have to do is put the objects in the water to see if they float (implying that this will be done during the lesson).	1,0/STR
	Now put the plasticine in the water and see if it floats.	1,0/SOL

12. Statements which call attention, even if the nature of the response is not clear or if there is no response, are coded SOL.

Example: Miss, Virginia is marking it out. 4,1/SOL

Substantive-Logical Meanings

In the Bellack system, the analysis of the substantive content of the discourse involved the use of two sets of categories, the first designed to code the subject matter under study and the second to code the cognitive processes involved in dealing with this subject matter. In the present study it was necessary to modify the analysis of substantive content in two ways. First, since Bellack's substantive meaning categories were subject-specific, and since the lessons under investigation were oriented more towards the processes involved in data gathering and processing than towards subject matter in the conventional sense, it was thought that no useful purpose could be served by developing substantive meaning categories in the Bellack sense. Second, because the cognitive processes used in the lessons were applied in the context of first hand experiences, it was found that the substantive-logical categories of the Bellack system were not directly applicable to much of the discourse.

In the present system, then, substantive meaning categories were disregarded and substantive-logical meaning categories modified to reflect the direct empirical nature of the lessons at hand. For this purpose, the science process model proposed by the American Association for the Advancement of Science (1965) in developing the program Science-A Process Approach yielded a conventient list of categories that have been found to be directly applicable to coding discourse related to empirical investigations in the classroom. The following list of categories, slightly modified from the AAAS version, has been used:

Observing
Classifying
Quantifying
Communicating
Predicting
Inferring
Hypothesizing
Defining Operationally
Interpreting Data
Controlling Variables
Experimenting
Forming Models

Because not all of the substantive discourse in the lessons could be regarded as being directly related to the specific empirical problem under investigation, it was necessary to retain a more broadly based set of categories, in addition to the empirical categories given above. The following list of categories, again slightly modified from Bellack, was used:

Fact Stating
Defining
Describing
Interpreting
Explaining
Evaluating
Opining
Justifying

The logical similiarities between the above two sets of categories is perhaps obvious. Nevertheless, the distinction between discussion of the specific problem of the lesson and discussion of more general substantive matters is of sufficient importance to justify the use of both sets.

With each substantive-logical statement is associated a "level of thought." on a three point scale as follows:

Level 0	Incorrect or implausible
Level 1	Correct
Level 2	Unique or unusual

The levels thus correspond to a rating scale, with the coder being required to make judgments on the scale value for each substantive-logical utterance. The judgments must be made with the grade level of the pupils in mind and in the context of the lesson at hand.

NOTE: At this point, the validity of the rating scale is under question, particularly since very few utterances are coded as level 2. Alternative procedures for the coding of levels of thought are being sought.

OBSERVING (OBS)

An observation is any reference to a specific object or event in the problem under study. An observation refers to something that is discerned by the senses (or by an instrument such as a magnifying glass which is a direct extension of the senses).

Examples:

1. This nail floats.
2. The small string sounds higher than the large one.
3. The balance arm is tipped to one side.

CLASSIFYING (CLS)

Classifying statements or actions are those that deal with the grouping or ordering of objects. Statements referring to the class to which an object belongs, an object which is an example of a class, the basis for forming a class, or the act of placing an object in a group or rearranging groups are examples of classification.

Examples

1. Nails are metal.
2. All these leaves are red.
3. Maybe we should group them by size.

QUANTIFYING (QUA)

Any reference to measurement, counting, arithmetic processes, spatial-temporal relationships, is considered to be quantifying. Serial ordering references are considered quantifying if it is clear that a unit of measurement is being used as the basis of ordering. Otherwise, serial ordering is considered to be classifying.

Examples

1. All these leaves look like triangles
2. Rock A displaces more water than rock B
3. It takes six washers (to balance the rock)

COMMUNICATING (COM)

Communicating is any reference to the description of objects or events in a manner that can be interpreted by other persons or in a manner which improves the efficiency of description. References to recording data, the use of symbols, tables, graphs, diagrams, etc., are part of this process. (Specific references to the use of such devices to interpret data, however, are part of the interpreting data process.)

Examples

1. We should draw a graph of that (However, a statement such as "the graph is a straight line is part of the process of interpreting data).
2. You can use the symbol ⊣⊢ to stand for a battery.
3. Record all your findings on that chart on the board.

PREDICTING (PRE)

In order to distinguish between predicting on the one hand and hypothesizing and inferring on the other, a relatively narrow interpretation of the term predicting is used. The term here refers to statements about specific future events or about extensions of the data at hand. Interpolating and extrapolating to find specific data values is included. However, references to the process of interpolation or extrapolation would be interpreting data.

Examples

1. What will happen to this rock if I drop it in the water?
2. If you make the string twice as long it will swing half as fast. (However, a statement such as "the longer the string, the slower it swings? is a hypothesis).
3. One small washer should balance it.

INFERRING (INF)

An inference is any statement about an event based only indirectly on observation. Conclusions drawn from inductive or deductive reasoning may be considered inferences. (see also hypothesizing and predicting).

Examples

1. I think the ball is hollow.
2. Nails sink because they are heavy.

HYPOTHESIZING (HYP)

A hypothesis is any expression of a relationship among two or more variables or any reference to an explanation for certain phenomena. A hypothesis is distinguished from a prediction in that the hypothesis must express some sort of generalization. Implicit in the statement of the hypothesis must be the idea that the hypothesis is to be tested. References to hypotheses that are predetermined (not formed by the speaker) are coded level 1 hypothesizing.

Examples

1. Today we are going to look at the relationship between the weight loss of a rock and the weight of the water it displaces. (Level 1)
2. Can anyone suggest one way in which we might make the pendulum swing faster?
3. What we have to test is whether all metals sink.

DEFINING OPERATIONALLY (OPD)

An operational definition is a statement in precise terms of the conditions necessary to identify an object or event.

Examples

1. Our units of weight will be big washers and small washers (assuming that the washers in question are on display).

2. A symmetrical object is one that can be folded to give matching halves
3. One liquid is more dense than another if the first liquid sinks when placed in the second

CONTROLLING VARIABLES (COV)

Statements dealing with the identification of the variables that might influence a particular experiemnt or of possible means of holding constant or otherwise deliberately manipulating these variables. The variables referred to should be other than those under direct investigation in the experiment at hand.

Examples

1. The strings have to be the same length.
2. Do you think it will matter if we do the experiment in the morning or afternoon?
3. If we are going to look at shape, all the pieces should be the same size.

INTERPRETING DATA (DAI)

Statements which indicate attempts to draw generalizations from gathered data. References to graphs, tables, maps, etc., other than those references defined under communicating are included here.

Examples

1. It looks like the time increases by two seconds each time.
2. From what we have done, is it correct to say that all metals sink?
3. The weight on each side of the balance is always the same when it is balanced.

EXPERIMENTING (EXP)

Since experimenting is a broad process which encompases all of the processes already identified, experimenting statements, as such, are rare and difficult to identify. For coding purposes, statements of trial and error procedures are coded experimenting, level 0. Other statements which directly refer to the conduct of an experiment can also be coded in this category.

Examples

1. Let's just try it with three washers (if the context makes clear that there is no particular reason why three rather than some other number is being tried or what the outcome might be).

2. We need to do an experiment to test that.
3. A different kind of experiment would be to slide a paper clip along the arm.

FORMULATING MODELS (MOD)

This process is included in order to complete the process model. However, because of the abstractness of the cognitive operations involved, it is not expected that this process will be observed with any great frequency in elementary science classes. Essentially, model formation statements are those which attempt to explain phenomena by analogy with something familiar or statements which are theoretical in nature (assuming that these statements refer directly to the phenomena under investigation in the lesson, otherwise product meanings are coded).

Examples

1. Electricity in wires is something like water flowing in pipes.
2. A magnet must be made up of many very small magnets because when a magnet is cut in half, each piece is also a magnet.

FACT STATING (FAC)

To state a fact is to give an account or report of a current, past, or future state of affairs or event. Statements of names, dates, places, etc., are examples of facts. In the context of scientific investigations statements of observed phenomena are facts but in the current system are coded as such only if the observation is not directly related to the lesson at hand.

Examples

1. What is the boiling point of water?
2. Can anyone remember whether a piece of plasticine sinks or floats?

DEFINING (DEF)

To define is to give the meaning of a term or to state the term associated with a given set of properties. Giving examples of objects in a particular class is also defining (except when the objects referred to are immediately at hand, in which case we use the process category CLS).

Examples

1. What do we mean when we say something is displaced:
2. Give me some examples of objects that float.

DESCRIBING (DES)

To describe an object or event is to list its properties or characteristics. The description may be in sense data terms or may be more quantitaive.

Examples

1. Can someone tell me what a pendulum looks like?
2. What are some of the parts of a bulb?

INTERPRETING (INT)

To interpret a statement is to express the statement in new words that are presumably more comprehensible than the original form. Interpreting is also used in the sense of interpreting data if the data is not related to the activity at hand.

Examples

1. What do we mean when we say that structure follows function?

EXPLAINING (EXP)

Explanations may broadly be regarded as attempts to answer the questions why? and how? or to give reasons for the occurrence of a particular phenomenon or the existence of some state of affairs. Explanations also refer to the effects of some event on some other event.

Examples

1. Why do steel ships float?
2. Because they displace water equal to the weight of the ship.
3. How can we balance a seesaw with two children of different weight?

EVALUATING (EVL)

An evaluative statement is one dealing with the worth or value of a particular statement or procedure. Evaluations are typically made by reference to some established criteria or standard.

Example

1. That is a poor experiment because it was conducted outdoors in the wind.

OPINING (OPN)

To give an opinion is to make a statement of a person's views about procedure or about the value fairness, worth, importance, or quality of an action, event, person, idea, plan or policy. In order for a statement to be an opinion, it must clearly represent the individual's personal view and must be open to the possibility of opposite views being tenable.

Examples

1. I don't think we should have to go outdoors.
2. I think it would be a good idea to take a field trip.

JUSTIFYING (JUS)

Justifying occurs when some reason is given for the holding of a particular opinion. Justification is distinguishable from evaluation in that evaluations deal with substantive matters while justifications deal with more personal matters. Thus, for example, a statement which criticizes a particular experimental procedure on the grounds of non control of variables would be coded as evaluation, whereas a personal view which is being defended would be coded justification.

Examples

1. A field trip would be nice because it gets us away from the classroom.

Coding Instructions For Substantive-Logical

1. Generally the decision to code process meaning or the other S/L categories is inferred from the context. At the beginning of the lesson process meanings are not coded until it becomes clear that the references are to specific data to be gathered in the lesson. At the end of the lesson, process meanings are used when the summary or discussion refers to the immediate data. Even if the general context is process, however, facts, definitions, etc., which are beyond the immediate data at hand are not coded or processed.

2. Statements which in themselves have no S/L meaning (i.e. rating reactions, control statements, etc. are coded in the S/L dimension as long as this can be inferred from the context. In such cases, no level is coded for S/L.

   Example: Put two more washers in. 4,5/SOL/QUA ...

3. Attempts at clarification, analogies, examples, etc. are coded as XPL if they have no reference to the immediate data.

4. The coding sheet gives the categories in order of priority. When more than one possible category applies, the one further down the list is coded.

5. Statements of the "try it and see" variety are coded as level 0, EXP.

6. Any explicit reference to data recording procedures is coded COM.

7. OBS is coded even if the statement refers to an observation made some time earlier. In this case, the level of the observation is omitted.

Example:	What happened to A in green	1,0/SOL/OBS/...
	It floated.	6,1/RES/OBS/...

10. A SOL which calls for an observation or fact is coded OBS or FAC but if the response adds to the observation or fact by giving an elaboration, explanation, or interpretation the higher cognitive level code is used.

Controlling

The notion of "control" in the classroom setting is fundamental to the study of teaching strategies. While the distribution of pedogogical moves gives a broad indication of degree of control by the teacher (or by a particular pupil in the small group setting), a more specific analysis of the nature of the controls exerted is needed if operational distinctions between teaching strategies are to be made. For example, it is useful to distinguish between the imperative and interrogative form of the soliciting move and to record instances of clarification, example giving, or calling attention to some event (a common occurence in science classes).

In identifying categories of control, a distinction is first made between statements, commands, and requests. A more detailed breakdown of the nature of these statements, commands, or requests is then given. A further aspect of control, that of rating, is treated as a separate dimension.

Controlling

GIV Gives or States
REQ Requests
COM Commands

→

Categories

TRE Teacher Requirement
PRE Pupil Requirement
PRF Performance
PRC Procedure
INT Statement of Intention
ATT Attention
RPT Repetition
ASS Assistance
CLA Clarification
ELA Elaboration
EXA Example

TEACHER REQUIREMENT (TRE)

A statement of teacher requirement is any statement that refers explicitly to what the teacher wants done in the lesson. Teacher statements of the variety "now what I want you to do is...", as well as pupil statements such as "teacher said we have to..." are examples. Teacher requirement is not used when the reference is incidental. Then a statement such as "now we have to..." does not clearly imply that it is the teacher that requires the act.

PUPIL REQUIREMENT (PRE)

Pupil requirement is the counterpart of teacher requirement. In its rare occurences, it takes the form, "John wants...", or "John said to..."

PERFORMANCE (PRF)

Performance refers to any command or request for immediate action by the receiver, other than those that can be coded by control categories.lower on the list. Performance is thus a sort of general control function. The specific nature of the act to be performed is coded under a type of reference.

Examples

Put another washer in.	.../COM/PRF/APP/...
Come here.	.../COM/PRF/ACP/...

PROCEDURE

Procedure statements differ from performance statements only in that the required action is not immediate. Procedure statements thus tend to occur more with STR and REA moves than with SOL moves, and with GIV rather than REQ or COM.

Examples

What we are going to do is see which liquids are heavier.	.../STR/.../GIV/PRC/ACT/...
This is how we weigh the rocks.	.../STR/.../GIV/PRC/APP/...

STATEMENT OF INTENTION (INT)

Statements of intention refer to activities that the speaker himself intends to perform or activities that he wishes the group to perform.

Examples

1. I'm going to use the plasticine.	.../GIV/INT/APP/...

2. Let's go get some water. .../GIV/INT/ACP/...

ATTENTION (ATT)

Attention is coded when the statement is intended to have the receiver change the focus of his attention. Such statements are evidenced by verbal conventions such as "look", "teacher", etc.

Examples

1. Teacher, Brian got his to work. .../REQ/ATT/APP/...
2. Look, it floats. .../COM/ATT/PHE/...

REPETITION (RPT)

Repetition is coded whenever the command or request is for the repetition of a statement or action or whenever a statement is voluntarily repeated. (The exception occurs when the repetition takes the form of a rating, in which case RPT is coded under rating rather than under controlling).

Examples

1. What did you say? .../REQ/RPT/ACV/...
2. Do that again. .../COM/RPT/ACP/...

ASSISTANCE (ASS)

Assistance is used in the narrow sense of explicit requests for help or statements indicating that help is to be given.

Examples

1. Teacher, can you show me how to do this. .../REQ/ASS/...
2. Here, let me hold the string. .../GIV/ASS/...

CLARIFICATION (CLA)

Clarification refers to a rewording or correction of a previous statement or action, to requests or commands for justification of a procedure or a substantive statement, to the giving of such justifications or repetitions, and to repetitions or requests for repetitions when it is obvious that the repetition stems from lack of clarity of previous actions.

Examples

1. Can you show us again how to slide the paper clip along the arm? .../REQ/CLA/...

2. What do you want the water for? .../REQ/CLA/...

ELABORATION (ELA)

Elaboration occurs typically in the context of clarification when new information is given or requested. More broadly, elaboration occurs when a point made in a previous statement is amplified or extended or when a response includes more information than required by the original solicitation.

Examples

1. Can anyone give me one variable that might influence how fast a pendulum swings? .../SOL/...

2. The length, because big grandfather clocks have pendulums that go slower than small clocks. .../RES/.../GIV/ELA/...

EXAMPLE (EXA)

Example is coded in the same context as elaboration and clarification when the elaboration or classification takes the form of a specific instance. Example is also coded for applications of phenomena or generalizations dealt with in the lesson. Example is not coded for simple ennumeration of instances of a phenomenom or event.

Examples

1. What are some of the uses of a pendulum? .../REQ/EXA/...

2. This table illustrates how your data should be recorded. .../GIV/EXA/...

Coding Instructions For Controlling

1. The basic distinction between COM, REQ, and GIV is whether the sentence is imperative, interrogative, or declarative, with the following exceptions:
   (a) Sentences using terms such as please, could you please..., would you... are considered to be requests even if it is not clear that a question is being asked. In this instance, overall tone of the sentence rather than the strict grammatical structure is considered.

(b) Statements of overall procedure are taken as declarative rather than imperative if the procedure is not to be immediately followed, or if the reference is in the form of "we have to..."

Example:	Today we have to look at how to use a paper clip to get a better idea of the weight of an object.	.../GIV/PRC/...

2. INT is coded only when the intention is to do something immediately. Otherwise PRC is more appropriate.

Examples:	I'm going to get me more water.	.../GIV/INT/...
	We should get more water.	.../GIV/PRC/...

3. When GIV/RPT is used, the entire coding of the previous sentence is repeated.

Example:	The green sinks.	4,5/DAG/REA/OBS/1/ / /PHE/
	What?	5,4/DAG/SOL/ / / /REQ/RPT/ACV/
	The green sinks.	4,5/DAG/REA/OBS/1/GIV/RPT/PHE/

Type of Reference

The establishing of a controlling dimension represents a departure in principle from the structure of the Bellack system in that the process of instruction for purposes of this study is conceptualized in terms of the controlling function. Nevertheless, when control is exercised, something must be controlled. Thus references are made to assignments, procedures, apparatus use, observed phenomena, and so on. Categories in this dimension are similiar to those in the "instructional" dimension in the Bellack system, with allowances being made for the direct experiences with phenomena that occur in science classes.

Statement
Logic reference
Language mechanics
Assignment
Procedure
Person
Action general
Action physical
Action vocal

Action cognitive
Action emotional
Apparatus
Individual results
Class results
Expected or desired outcome
Event or phenomenon
Recording equipment

STATEMENT (STA)*

Statement is coded when the reference is to a previous verbal utterance, particularly to the meaning, truth, validity or propriety of that utterance. Statement typically occurs with a rating. (see Rating).

Examples

Good point	.../STA/POS/
Wrong.	.../STA/NEG/

LOGIC REFERENCE (LOG)

A logical reference is any explicit reference to logical processes such as defining, explaining, inductive or deductive reasoning, etc.

Examples

There's a fallacy in that argument.	.../LOG/NEG/
Is it possible to prove a hypothesis to be correct?	.../LOG/...

LANGUAGE MECHANICS (LAM)

This category refers to explicit reference to language usage, grammar, or spelling.

Examples

1. P. I want s some water.
   T. I want some water.
2. How do you spell pendulum?

ASSIGNMENT (ASG)

Assignment is coded when reference is to homework, projects, reports, etc.

---

* Definitions of those categories drawn from the Bellack system follow Bellack (1966), Chapter 2, with only minor modifications and with example changes.

outside the immediate tasks being performed during the lesson.

Examples

1. Before next week's class, we have to collect some bottles.
2. Do we have to write this up afterwards?

PROCEDURE (PRC)

This category is the same as PRC in the controlling dimension but is used when procedures are referred to but the controlling category is clarification, elaboration, teacher requirements, etc. Thus, a coding sequence such as GIV/PRC/PRC/ is redundant, whereas GIV/CLA/PRC/ is meaningful.

Examples

1. What we are going to do is see if each object floats or sinks. .../STR/.../GIV/PRC/...
2. In other words, what we have to do is take each object in turn and drop it into the water. .../STR/.../GIV/CLA/PRC/...

PERSON (PER)

Person is used when the reference is to an individual's physiognomy, dress, appearance, or expression or when personal experience is being discussed.

Examples

1. Roll up your sleeve. .../COM/PRF/PER/...
2. We have a pendulum clock at home. .../GIV/EXA/PER/

ACTIONS (ACT, ACP, ACV, ACC, ACE)

The specific nature of the performance, action, or event referred to in the controlling dimension is coded under one of the action headings if none of the more specific codes lower in the list apply. These categories are most easily defined through examples.

Examples*

What are you doing? (The nature of the act is not clear.) ACT

---

* Additional examples are given in Bellack (1966). Bellack's examples are applicable here even though the substantive nature of the lesson is quite different.

Go up and get some paper towel. (Reference to physical movement.)	ACP
What did you say?	ACV
I'll have to think about that.	ACC
That's not fair.	ACE

APPARATUS (APP)

Specific references to apparatus or materials, procedures for apparatus use, apparatus manipulation, etc. are coded in this category. In particular, APP overrides the action categories defined above when the action refers to action with apparatus.

Examples

1.	Put the nail in another notch.	.../COM/PRF/APP/...
2.	We need more batteries.	.../GIV/PRC/APP/...

INDIVIDUAL RESULTS (IND)

This category refers to any statement in which the results obtained by a specific student or group are discussed. Typical instances are when the teacher asks the group to describe the outcome of an experiment or the activity, or when a student reports such outcome to the teacher or to the class.

Examples

1. What object did you get to float? (Teacher asking a particular group).
2. We got six bulbs to light.

CLASS RESULTS (CLA)

This is the counterpart of individual results when the reference is to the class as a whole. A teacher request for results which is directed at the entire class rather than an individual, a teacher summary of reported individual results, or references to data compiled by the entire class are included. Also included are simultaneous responses by a number of pupils to the requests for results. However, if an individual responds with his own results, IND is coded even if the teacher does not address the request to any particular individual.

Examples

1. It looks as if some people got it to float and others got it to sink, .../CLA/...
2. What object did you get to float? (Teacher addresses no particular individual or group). .../CLA/...
3. Rubber, wood, bottle caps. (Response by one student to the request in Example 2), .../IND/...

EXPECTED OR DESIRED OUTCOME (EXP)

This category is used in preference to IND or CLA when it is clear that what is being referred to is an anticipated result or a result that had been identified earlier as desirable.

Examples

1. How many people found that the pendulum went slower as you made it longer?
2. You were supposed to put the liquids in order of density.

EVENT OR PHENOMENON (PHE)

PHE is coded whenever the reference is to an observation, a predicted or inferred event, or the reporting of a previous observation.

Examples

1. It balances.
2. I think it will sink. (In this case PHE overrides ACC, the other possible coding).

RECORDING EQUIPMENT (REC)

This category is used to code any reference to the presence of the recording equipment or personnel. The category overrides other types of reference when it does occur since it indicates possible interference in the operation of the class by the act of recording.

Examples

1. We're going to be on TV.

2. What's this speaker for?

3. Can you help me with this string? (Student to operator). .../SOL/.../REQ/ASS/REC/...

Coding Instructions For Type of Reference

1. As in other dimensions, categories lower in the list override those higher when more than one code is possible for a single sentence. In particular, APP overrides ACP in the frequent instance of carrying out a physical action using apparatus.

2. ACE is used for emotional response or reactions to previous statements even if some other response is indicated by that statement.

   Example: T. Go outside and get some snow. .../SOL/.../COM/PRF/APP/...

   P. I don't think that's fair. .../REA/.../ACE/NEG/...

   This example also illustrates the use of REA rather than RES following a SOL since the pupil statement is not the type that would normally be elicited by the original SOL.

3. Ordinary references to past observations are coded PHE rather than IND. The exception is when a pupil is called upon by the teacher or by another pupil to report what he has observed.

   Examples: Did the string stretch? .../SOL/OBS/ /REQ/PRF/PHE/

   No. .../RES/OBS/ /GIV/PRF/PHE/

   Note: The level of observation is omitted here to indicate that a past observation is being reported.

4. If a SOL calls for a particular action such as ACP and the response is with a different form of action, the SOL is coded according to the expected response and the RES according to the actual response. If the RES contains two types of response, the one elicited by the SOL is recorded.

   Examples: T. Try making the string shorter. SOL/.../COM/PRF/APP

   1. I already did that. REA/.../GIV/PRF/STA/POS/

   2. Okay (while actually making the string shorter) RES/.../GIV/PRF/APP/POS/

3. I don't think that's a good idea.	REA/.../GIV/PRF/ACC/NEG/

5. Use ACP for acts of writing, looking, listening, etc.

Examples:		
	1. Listen to this.	.../SOL/.../COM/PRF/ACP/...
	2. Look. (If the command "look" was made with reference to an observation, PHE would be used rather than ACP).	.../SOL/.../COM/ATT/ACP/...

6. ACV is coded only when reference is specifically to a vocal action and not merely when a SOL calls for a verbal response.

Examples:		
	What did you say?	.../SOL/.../REQ/REP/ACV/
	I told you that before.	.../REA/.../GIV/PRF/ACV/...

Ratings

Ratings are evaluative judgments about previous statements or actions. As such, ratings serve as control devices in much the same way as the requests, commands and statements of various types referred to in the controlling dimension. That is, judgments about particular behavior are likely to have some influence on whether or not such behavior is continued. Psychologically, of course, ratings are related to the concept of reinforcement.

Some category systems, notably that of Flanders, contain only two rating categories, namely praise and criticism. The Bellack system, in contrast, contains no less than six such categories on a scale from positive to negative as well as two other categories dealing with requests for ratings. In our adaptation of the Bellack approach to ratings, minor modifications have beem made because it was found that two categories, admitting and not admitting could not reliably be distinguished from positive and negative. The rating categories that remain are as follows:

Positive
Qualifying
Repeating
Negative
Positive/ negative

The definitions of these rating categories are straightforward and are illustrated here by examples. Generally, ratings occur for reaction moves only. However, ratings are assigned to responses when the original solicit move calls for a yes or no response. The PON move is always used with SOL.

POSITIVE (POS)

Examples

1. Right.
2. OK.
3. Good idea.

QUALIFYING (QAL)

Examples

1. Right, but you mean shape.
2. That's true, except that...

REPEATING (RPT)

Examples

1. P. It has sides.

   T. Right, it has sides. .../RPT/...
2. Jackie says that the balance tips.

NEGATIVE (NEG)

Examples

1. No.
2. Wrong.
3. I don't agree.

POSITIVE/ NEGATIVE (PON)

This category refers to a request for a rating. It is thus always used with a SOL rather than a REA or RES.

Examples

1. Is this right?

2. How does this graph look?

Coding Instructions For Rating

1. If a particular statement contains both a rating and some other substantive and/or instructional content, REA is coded for the move, with the appropriate rating. The remainder of the statement is coded as if the rating did not exist.

   Example: Right, except that we need to use small washers to make up the difference. .../REA/QUA/GIV/PRC/APP/QAL/...

2. NEG is used with SOL when a command is given in the form of a prohibition of some action.

   Example: Don't touch it. .../SOL/.../COM/PRF/ACP/NEG/...

   However, such commands are coded REA rather than SOL when it is clear that the prohibition was elicited by some attempt on the part of the receiver to engage in the prohibited action.

3. Rating reactions ordinarily occur with reference to a previous statement. However, the code STA/NEG is used only when no higher category in type of reference is found to apply. Thus, for instance, if the statement being rated refers to apparatus, the proper coding is APP/NEG.

   Examples:

   1. P. This plasticine is wet.

      T. You should not get it wet. .../REA/.../APP/NEG/...

   2. P. It sinks

      T. You mean it floats. .../REA/.../PHE/NEG/...

   3. P. A pendulum should go faster as the string is made longer. .../RES/HYP/1/

      T. Good. .../REA/HYP/.../STA/POS/...

4. Although ratings must be given (declared) or requested, GIV or REQ is not coded for a rating reaction since this would be redundant. (See above example).

Physical Action

It can be argued that, in the science laboratory setting, nonverbal actions, particularly apparatus manipulations, form an essential part of the ongoing behaviour. Since few existing interaction analysis systems contain catagories for nonverbal action, few guides to the development of such catagories exist. At first it was thought that the literature in the area of time and motion study as used in job analysis would be of some value. It was found, however, that job analysis systems tended to concentrate on such acts as reaching and grasping which, while no doubt useful in improving the efficiency of performance of job tasks, were regarded as rather trivial for purposes of this study. The physical action catagories given below were therefore devised by viewing the videotapes and listing those actions that appeared to contribute to the analysis of the structure of the lessons. These physical action catagories must be regarded as rather tentative at this point since work is continuing in this area. Also, it should be noted that the list is by no means exhaustive in that it excludes, notably, gestures, facial expressions, and related nonverbal elements of the process of communication.

Apparatus Manipulation

Setting Up
Dismantling
Adjusting
Taking Reading or Measurement
Molding, Shaping, Cutting
Watching
Nonspecific Action
No Action

Recording

Writing
Drawing
Graphing
Tabulating
Calculating
Non-specific recording

Management Movement

Fetching/ returning
Changing position
Raising hand
Off task
Clean up

Recording equipment
Non-specific management
Not codable

Most of these categories are self explanatory. It has been found that physical actions can be coded directly from the vodeotape in a straightforward manner. The NON, NOR, and NOM categories have been included because very often it is possible to observe that an individual is, in fact, manipulating apparatus or writing or moving about, but the specific nature of the movement cannot be detected from the videotape. Particularly in recording, it is often not possible to obtain a sufficiently close view to determine which of the specific recording acts is being carried out, even though it is quite obvious that the individual is writing something. Similiarly, many seemingly arbitrary movements of apparatus are made. The NON, NOR, and NOM categories thus identify the action only as apparatus manipulation, recording, or management.

Coding Instructions For Physical Action

1. Physical actions are coded simultaneously for three persons--the teacher and the pair of pupils under analysis. Because of the recording procedure used, the teacher is very often not visible. In this situation, NOC is used.

2. NOC refers specifically to a person not being visible on the recording. However, if the view of a person is obstructed only momentarily, and if the person is observed to be engaged in the same action both before and after the obstruction, the type of action coded before and after is inferred to have occurred during the time of the obstruction.

3. Verbal action can often give a clue to physical action in cases where the nonspecific code might otherwise have to be used.

Example: P1. Write that
in the table.

P2. (Writes)

It is inferred here that P2 is, indeed, writing in the table. TAB is therefore coded.

Interpersonal

The preliminary version of the system included a set of catagories for the coding of social-emotional tone of the discourse, adapted from Bales (1950) and Barker (1963). However, it has been found that too few strictly personal exchanges occur to establish the reliability of the proposed catagories. In the rare instances of personal exchanges, the discourse is coded off task in the lesson phase dimension but can be coded in the move, controlling and rating dimensions in the usual way. Personal exchanges can be distinguished from other off task behaviours by the coding of PER in the type of reference dimension.

CODING PROCEDURES AND RELIABILITY TESTS

Following Smith and Meux (1963) and Bellack (1966), a team coding procedure has been adopted whereby each transcipt is coded by a pair of coders. Originally the two coders worked together, developing a consensus on each sentence as they worked through the transcript. However, an analysis of transcript segment coded by two separate teams revealed that even a pair of coders can develop a "set" in which the adoption of a particular coding convention caused distortion in the coding. A major problem thus became the identification of "clusters" of coding differences caused by a single difference in interpreting the meaning of a catagory.

The procedure now being followed is for each coder to first independently code the transcript. A set of matrices are then developed which permit the identification of areas of agreement and disagreement in each dimension. A sample comparison matrix for the controlling dimension is shown in Figure 1. It is usually found that coding differences occur in one or two major clusters in each dimension, each cluster being caused by a single difference in interpretation. The computer program that produces the comparison matrices also gives a listing of sentences on which disagreements occur for each dimension. The sentence listing is used as a basis for discussion by the entire group of

Coder A

Coder B	STA	LOG	LAM	ASG	PRC	PER	ACT	ACP	ACV	ACC	ACE	APP	IND	CLA	EXP	PHE	REC	Blank
STA	5				1			2				9						
LOG																		
LAM			1									2	1					
ASG																		
PRC					1			1				1						
PER						2		3										
ACT	2				2	1	1	2				1						
ACP							1	4										
ACV							7		9	1		2	1			1		3
ACC										6		2						
ACE											1		1					
APP	5				1		2	29				64		1		1		
IND												14	43	10		3		
CLA												2	4	9		4		
EXP	1				1		1	9		3		11				3		
PHE	8							7		3		5				30		3
REC																6	15	
Blank	5				1		6	8	3	1	2	23	1	1	3	15	2	12

Significant Coder Disagreements

1. ACP - APP
2. APP - Blank
3. PHE - Blank
4. APP - IND

Initial Agreement Index $R_1 = 0.44$

Final Agreement Index $R_2 =$

Figure 1 Sample Comparison Matrix "Type of Reference" dimension

four coders of clusters in which there are more than ten disagreements between the pair. It has been found that a single source of error usually accounts for most of the differences in a cluster. Therefore, once agreement has been reached on a single point, it is possible to correct the coding for a relatively large number of sentences.

By the time coding is complete for the initial tapes to be used in the descriptive analysis, it is anticipated that a clear picture will emerge of any areas in which persistent disagreements occur. This will permit an estimate to be made of the degree of confidence with which certain interpretive statements can be made in the overall description of the classes. Also, at this stage, it should be possible to refine the system somewhat if necessary for use with the new data.

Table 1 gives a sample set of coefficients of agreement between two raters for a transcript consisting of some 450 discourse units. The major difference between the R_1 and R_2 coefficient reflects the degree of consensus possible when all raters discuss the specific areas of disagreements that appear in the matrix analysis.

A more formal reliability test is being conducted and will be reported in later descriptions of the system. This test involves proportion of agreements between two pairs of coders after each pair has followed the procedure already described. A new indices of agreement will be more generally valid than the working indices already described since they will indicate differences between pairs rather than between individuals. It must be noted, however, that the reported indices will be based on pairs who have interacted with each other throughout the study. Because of the length of the training procedure, it will not be possible to determine the degree of agreement between independent pairs who have only the catagory definitions and the coding instructions in common.

This latter point raises certain more general questions concerning the usefulness of indices of inter rater reliability. Certainly it can be argued that if obtaining a high degree of agreement in coding becomes the overriding consideration, then there is a danger of reducing the system to a

DIMENSION	Initial Proportion Agreement R_1	Final Proportion Agreement R_2
Move		
Lesson Phase		
Substantive-logical		
Controlling		
Sentence form		
Control type		
Type of reference		
Ratings		

Table 1

Sample, Inter-Rater Agreement Indices before and after discussion of Matrix.

set of relatively trivial but easily detected categories. This is especially true if the aim is to produce a system which leads to highly consistent coding by observers not connected with the original study.

If the coding system is interpreted simply as a tool to assist investigators in describing what happens in classrooms, then a case can be made that as long as an investigator arrives at statements about the structure of the classroom that can be verified by other investigators, possibly using entirely different coding systems, then the coding system has served a useful purpose. Reliability then becomes an issue of consistency in the results rather than of agreement in the coding of specific categories. A case in point is that of Piaget. In the Piaget studies, no emphasis was placed on developing indices which would indicate whether different observers would rate children's responses in exactly the same way. Rather, the concern was with whether consistent interpretations could be made of a variety of differently worded responses. In spite of its apparent subjectivity, Piaget's method has been most fruitful in formulating and testing a theory of intellectual development.

A similiar issue is raised by Herbert and Attridge (1974). These authors argue that high inter rater agreement in coding can not be interpreted. as meaning high reliability of scores developed from the categories. That is, the psychometric meaning of reliability, that of degree to which an observed score is indicative of a hypothetical "true score", is not conveyed by an inter rater agreement index. Unfortunately, Herbert and Attridge do not suggest any more direct methods of approaching the reliability problem.

The position taken in this study is that high inter rater agreement is essential if transcripts coded by different raters are to be pooled for purposes of analysis. Thus, within a single study, the data must be coded in a consistent manner. High agreement on the part of raters outside the original group can be interpreted as indicating precision in the operational definitions of categories and in coding conventions. Such agreement may not be crucial if new raters apply the system to new data and then interpret the data in a manner consistent with their understanding of the categories. A more genuine indication of reliability in the psycometric sense can be obtained only if classes with similar characteristics are coded in a similar way by different raters. Unfortunately, it would be impractical in most situations to establish a reliability study of this nature.

APPENDIX C

SUMMARIES OF REGRESSION ANALYSES FOR CLASSROOM BEHAVIORS: EXPERIMENTAL STUDY

Table 1

Summary of Regression Analysis: Ratio 1 (No Talk/Total Time)

		Round 1				Round 2			
		Segment							
Effect		1	2	3	4	1	2	3	4
Teachers	R^2	.196	.061	.110		.045	.370	.325	
	ΔR^2	.162	.047	.059		.007	.350	.299	
	dF	3,94	3,108	3,95		3,61	3,92	3,66	
	F	6.31	1.80	2.09		.14	17.03	9.76	
Treatment x Teachers	R^2	.290	.199	.213		.156	.447	.394	
	ΔR^2	.094	.138	.103		.111	.077	.069	
	dF	3,91	3,105	3,92		3,58	3,89	3,63	
	F	4.01	6.03	4.01		2.54	4.15	2.39	
ATI's	R^2	.395	.262	.264		.373	.579	.565	
	ΔR^2	.105	.063	.051		.217	.132	.171	
	dF	13,78	12,93	12,80		14,44	12,77	13,50	
	F	4.52	.66	.62		1.08	2.01	1.51	
Treatment	R^2	.395	.202	.204		.373	.379	.565	
	ΔR^2	<.001	<.001	<.001		<.001	<.001	<.001	
	dF	1,77	1,92	1,79		1,43	1,76	1,49	
	F								

Table 2

Summary of Regression Analysis:

Ratio 2 (Listening/Total Time)

		Round 1				Round 2			
					Segment				
Effect		1	2	3	4	1	2	3	4
Teachers	R^2	.156	.033	.108		.073	.175	.182	
	ΔR^2	.150	.024	.063		.060	.160	.178	
	dF	3.94	3,108	3,95		3,61	3,92	3,66	
	F	5.57	.89	2.27		1.31	5.95	4.79	
Treatment x Teachers	R^2	.230	.243	.279		.217	.273	.185	
	ΔR^2	.074	.210	.171		.144	.098	.003	
	dF	3,91	3,105	3,92		3,58	3,89	3,63	
	F	2.91	9.71	7.28		3.56	4.00	.07	
ATI's	R^2	.348	.302	.362		.467	.476	.352	
	ΔR^2	.118	.059	.083		.250	.203	.167	
	dF	14,77	13,93	11,81		13,45	14,75	13,50	
	F	.99	.60	.95		1.62	2.07	.99	
Treatment	R^2	.348	.302	.362		.495	.496	.352	
	ΔR^2	<.001	<.001	<.001		.028	.020	<.001	
	dF	1,76	1,92	1,80		1,44	1,74	1,49	
	F								

Table 3

Summary of Regression Analysis:

Ratio 3 (P. Action/Total Time)

		Round 1				Round 2			
		Segment							
Effect		1	2	3	4	1	2	3	4
Teachers	R^2	.022	.072	.054		.109	.248	.329	
	ΔR^2	.015	.071	.054		.108	.247	.317	
	dF	3,94	3,108	3,95		3,61	3.92	3,66	
	F	.48	2.75	1.81		2.46	10.09	10.40	
Treatment x Teachers	R^2	.059	.199	.167		.161	.269	.394	
	ΔR^2	.037	.127	.113		.052	.021	.065	
	dF	3,91	3,105	3,92		3,58	3,89	3,63	
	F	1.19	5.55	4.17		1.19	.84	2.25	
ATI's	R^2	.150	.256	.211		.470	.338	.626	
	ΔR^2	.091	.057	.044		.309	.069	.232	
	dF	12,79	12,93	12,80		13,45	12,77	13,50	
	F	.70	.79	.37		2.01	.66	2.38	
Treatment	R^2	.150	.256	.211		.522	.352	.626	
	ΔR^2	<.001	<.001	<.001		.052	.014	<.001	
	dF	1,78	1,92	1,79		1,44	1,76	1,49	
	F								

Table 4

Summary of Regression Analysis:

Ratio 9 (Pupil Structuring)

		Round 1			Round 2		
		Segment					
Effect		1	2	3	1	2	3
Teachers	R^2	.045	.185	.180	.072	.036	.062
	ΔR^2	.024	.149	.145	.039	.036	.051
	dF	3,88	3,100	3,89	3,59	3,86	3,61
	F	.73	6.10	5.24	.82	1.06	1.10
Treatment x Teachers	R^2	.063	.238	.191	.097	.050	.157
	ΔR^2	.018	.053	.011	.025	.014	.095
	dF	3,85	3,97	3,86	3,56	3,83	3,58
	F	.54	2.25	.38	.51	.40	2.18
ATI's	R^2	.130	.429	.286	.259	.260	.363
	ΔR^2	.067	.191	.095	11,45	12,71	10,48
	dF	12,73	13,84	13,73	11,45	12,71	10,48
	F	.46	2.16	.74	.89	1.67	1.55
Treatment	R^2	.130	.428	.286	.259	.267	.363
	ΔR^2	<.001	<.001	<.001	<.001	.007	<.001
	dF	1,72	1,83	1,72	1,44	1,70	1,47
	F						

Table 5

Summary of Regression Analysis: Ratio 10 (Pupil Soliciting)

		Round 1			Round 2		
		Segment					
Effect		1	2	3	1	2	3
Teachers	R^2	.087	.027	.066	.089	.056	.076
	ΔR^2	.087	.026	.049	.089	.052	.042
	dF	3,88	3,100	3,89	3,59	3,86	3,61
	F	2.80	.89	1.55	1,92	1.57	.92
Treatment x Teachers	R^2	.156	.046	.093	.245	.131	.113
	ΔR^2	.069	.019	.027	.156	.075	.037
	dF	3,85	3,97	3,86	3,56	3,83	3,58
	F	2.31	.64	.85	3.86	2.39	.80
ATI's	R^2	.309	.129	.258	.486	.233	.419
	ΔR^2	.153	.083	.165	.241	.102	.306
	dF	14,71	13,84	14,72	11,45	14,69	13,45
	F	1.00	.61	1.00	1.91	.65	1.82
Treatment	R^2	.309	.129	.258	.493	.266	.419
	ΔR^2	<.001	<.001	<.001	.007	.033	<.001
	dF	1,70	1,83	1,71	1,44	1,68	1,44
	F						

Table 6

Summary of Regression Analysis:

Ratio 11 (Pupil Responding)

		Round 1			Round 2		
		Segment					
Effect		1	2	3	1	2	3
Teachers	R^2	.092	.020	.131	.105	.013	.067
	ΔR^2	.084	.020	.127	.103	.012	.052
	dF	3,88	3,100	3,89	3,59	3,86	3,61
	F	2.71	.68	4.34	2.26	.34	1.13
Treatment x Teachers	R^2	.132	.066	.136	.154	.027	.123
	ΔR^2	.040	.046	.005	.049	.014	.056
	dF	3,85	3,97	3,86	3,56	3,83	3,58
	F	1.30	1.59	.16	1.08	.39	1.23
ATI's	R^2	.338	.197	.293	.433	.339	.264
	ΔR^2	.206	.131	.157	.279	.312	.141
	dF	14,71	14,83	13,73	11,45	12,71	14,44
	F	1.57	.61	1.24	2.01	2.79	.60
Treatment	R^2	.338	.197	.293	.482	.339	.264
	ΔR^2	<.001	<.001	<.001	.049	<.001	<.001
	dF	1,70	1,82	1,72	1,44	1,70	1,43

Table 7

Summary of Regression Analysis:

Ratio 12 (Pupil Reacting)

		Round 1			Round 2		
		Segment					
Effect		1	2	3	1	2	3
Teachers	R^2	.167	.033	.026	.053	.017	.133
	ΔR^2	.124	.024	.019	.038	.016	.112
	dF	3,94	3,108	3,95	3,61	3,92	3,66
	F	4.67	.89	.61	.81	.50	2.84
Treatment x Teachers	R^2	.184	.082	.046	.070	.048	.180
	ΔR^2	.017	.049	.020	.017	.031	.047
	dF	3,91	3,105	3,92	3,58	3,89	3,63
	F	.63	1.86	.64	.35	.96	1.20
ATI's	R^2	.269	.277	.179	.265	.183	.412
	ΔR^2	.085	.195	.133	.195	.135	.232
	dF	12,79	13,92	13,79	13,45	12,77	14,49
	F	.76	1.91	.98	.91	1.06	1.38
Treatment	R^2	.269	.277	.179	.215	.216	.412
	ΔR^2	<.001	<.001	<.001	<.001	.033	<.001
	dF	1,78	1,91	1,78	1,44	1,76	1,48
	F						

Table 8

Summary of Regression Analysis:

Ratio 14 (QUA/Total SIL)

		Round 1			Round 2		
		Segment					
Effect		1	2	3	1	2	3
Teachers	R^2	.136	.092	.002	.207	.234	.075
	ΔR^2	.134	.078	.001	.207	.214	.047
	dF	3,94	3,108	3,95	3,61	3,92	3,66
	F						
Treatment x Teachers	R^2	.163	.155	.138	.242	.281	.110
	ΔR^2	.027	.063	.136	.035	.047	.035
	dF	3,91	3,105	3,92	3,58	3,89	3,63
	F						
ATI's	R^2	.238	.356	.242	.368	.339	.213
	ΔR^2	.075	.201	.104	.126	.058	.113
	dF	10,81	14,91	12,80	12,46	12,77	14,49
	F						
Treatment	R^2	.238	.356	.242	.375	.344	.213
	ΔR^2	<.001	<.001	<.001	.007	.005	<.001
	dF	1,80	1,90	1,79	1,45	1,76	1,48
	F						

Table 9

Summary of Regression Analysis

Ratio 25 (Proportion of Apparatus References)

		Round 1			Round 2		
		Segment					
Effect		1	2	3	1	2	3
Teachers	R^2	.011	.032	.022	.039	.288	.036
	ΔR^2	.008	.009	.016	.036	.256	.031
	dF	3,94	3,108	3,95	3,61	3,92	3,66
	F						
Treatment x Teachers	R^2	.051	.068	.082	.197	.300	.103
	ΔR^2	.040	.036	.060	.158	.012	.067
	dF	3,91	3,105	3,92	3,58	3,89	3,63
	F						
ATI's	R^2	.736	.140	.312	.461	.388	.303
	ΔR^2	.685	.072	.230	.264	.088	.200
	dF	12,79	13,92	14,78	13,45	13,76	14,49
	F						
Treatment	R^2	.736	.140	.312	.462	.424	.303
	ΔR^2	<.001	<.001	<.001	<.001	.036	<.001
	dF	1,78	1,91	1,77	1,44	1,75	1,48
	F						

Table 10

Summary of Regression Analysis:

Ratio 26 (Proportion of Individual Results)

		Round 1			Round 2		
			Segment				
Effect		1	2	3	1	2	3
Teachers	R^2	.118	.072	.051	.230	.040	.023
	ΔR^2	.103	.071	.039	.216	.039	.021
	dF	3,94	3,108	3,95	3,61	3,92	3,66
	F						
Treatment x Teachers	R^2	.168	.076	.055	.354	.138	.094
	ΔR^2	.050	.004	.004	.124	.098	.071
	dF	3,91	3,105	3,92	3,58	3,89	3,63
	F						
ATI's	R^2	.280	.138	.108	.466	.298	.209
	ΔR^2	.112	.062	.053	.112	.160	.115
	dF	14,77	13,92	10,82	12,46	14,85	13,50
	F						
Treatment	R^2	.280	138	.108	.496	.304	.209
	ΔR^2	<.001	<.001	<.001	.030	.006	<.001
	dF	1,76	1,91	1,81	1,45	1,84	1,49
	F						

REFERENCES

REFERENCES

Abraham, M.R. "The effect of grouping on verbal interaction during science inquiries." Journal of Research in Science Teaching, 13, 127-136, 1976.

American Association for the Advancement of Science. An Evaluation Model and its Application. Washington: AAAS, 1968.

Aschner, M.J. "The Analysis of Verbal Interaction in the Classroom." In Arno Bellack (Ed.) Theory and Research in Teaching. New York, Bureau of Publications, Teachers College, Columbia University, pp. 53-78, 1961.

Bales, R.F. Interaction Process Analysis. Cambridge, Mass: Addison-Wesley, 1950.

Bales, R.F. Personality and Interpersonal Behaviour. New York: Holt, Rinehart and Winston, 1970.

Barker, R.G. Ecological Psychology. Stanford California, Stanford University Press, 1963.

Bellack, A.A., et. al. The Language of the Classroom. New York: Teachers College Press (Columbia University) 1966.

Berliner, D.C. and Cahen, L.S. Trait-treatment interaction and learning. In F.N. Kerlinger (Ed.) Review of Research on Education, Vol. 1, Itascoe, Ill., Peacock, pp. 58-94, 1973.

Bernstein, B. Class, codes and control. London: Routledge and Kegan Paul, 1971.

Blishen, B.R. A Socio-economic Index for Occupations in Canada. Canadian Review of Sociology and Anthropology, 4:41-53, February, 1967.

Bracht, G.H. Experimental factors related to aptitude-treatment interactions. Review of Educational Research, 40, 627-646, 1970.

Brookover, W.B., Peterson, A., Thomas, S. Self-Concept of Ability and School Achievement, Office of Research and Publications, College of Education, Michigan State University, Michigan, 1962.

Brophy, J. and Evertson, C. Process-product correlations in the Texas Teacher Effectiveness Study: Student Sex, grade and socioeconomic status differences in classroom process measures. Report No. 74-4. Austin, Texas: Research and Development Centre for Teacher Education, University of Texas, 1974.

Chiappeta, E.L. A Review of Piagetian Studies Relevant to Science Instruction at the Secondary and College Level. *Science Education*, Vol. 60, No. 2, 253-262, 1976.

Crandell, V.C., Katkovsky, W., Crandall, V.J. Children's Beliefs in their Own Control of Reinforcements in Intellectual-Academic Achievement Situations. *Child Development*, 36, 91-109, 1965.

Crocker, R.K. *Elementary Science Curriculum Study*. (Teaching Guide, Volumes I and II). Toronto: McGraw-Hill Ryerson, 1973.

Crocker, R.K., Bartlett, K.R. and Elliott, H.G. A Comparison of Structured and Unstructured Modes of Teaching Elementary Science Process Activities. *Journal of Research in Science Teaching*, 13, 267-274, 1976.

Cronbach, L.J. Research on Classrooms and Schools: Formulation of Questions, Design and Analysis. Stanford, California, Stanford Evaluation Consortium, 1976.

Cronbach, L.J. and Snow, R.E. *Aptitudes and instructional methods*. New York: Irvington Publishers, 1977.

Cusick, P.A. *Inside the High School*; the student's world. New York, Holt, Rinehart and Winston, 1973.

Dreeben, R. *The Nature of Teaching: Schools and the Work of Teachers*. New York: Scott Foresman, 1970.

Dunkin, M.J. and Biddle, B.J. *The Study of Teaching*. New York, Holt, Rinehart, and Winston, 1974.

Eysenck, B.G. *Junior Eysenck Personality Inventory*. Education and Industrial Testing Service, California, 1963.

Flanders, N.A., Anderson, J.P. Amidon, E.J. Measuring Dependence Proneness in the Classroom. *Educational and Psychological Measurement*, Vol. XXI, No. 3, 1961.

Flanders, N.A. Analyzing Teacher Behaviour. Reading Mass: Addison Wesley, 1970.

Gates, A.I., MacGintie, W.H. Gates-MacGintie Reading Tests. Teachers College, Columbia University, New York, 1965.

Griffiths, A.K. and Crocker, R.K. Achievement in Individualized and Conventional Chemistry Courses and its Interactions with Selected Academic and Personality Variables. The Alberta Journal of Educational Research, Vol. XXII, No. 2, June, 1976.

Grannis, J.G. Columbia Classroom Environment Project. Extracts from final Report, 1973 (Mimeo).

Hunt, D.E. Matching Models in Education. Monograph Series No. 10. Ontario Institute for Studies in Education, 1971.

Hunt, D.E. Person-environment interaction: A challenge found wanting before it was tried. Review of Educational Research, 45, 209-230, 1975.

Hunt, D.E. and Sullivan, E.V. Between Psychology and Education. Hinsdall, Illinois, Dryden Press, 1974.

Hurd, P.D. and Rowe, M.B. "A Study of Small Group Dynamics and Productivity in the BSCS Laboratory Block Programme". Journal of Research in Science Teaching, 4:67-73, 1966.

Inhelder, B, and Piaget, J. The Growth of Logical Thinking from Childhood to Adolescence. New York: Basic Books, 1958.

Jackson, P.W. Life in Classrooms. New York: Holt, Rinehart and Winston, 1968.

King, Ethel M. (Ed.) Manual for Canadian Tests of Basic Skills. Thomas Nelson & Sons Ltd. (1968).

Gage, N.L. Four Cheers for Research on Teaching. (Unpublished Paper, Stanford University, 1976).

Kuhn, T. The Structure of Scientific Revolutions. (Second Edition). Chicago: University of Chicago Press, 1970.

Kounin, J.S. Discipline and Group Management in Classrooms. New York: Holt, Rinehart and Winston, 1970.

Lundgren, U.P. Frame Factors and the Teaching Process. Stockholm: Almquist and Wiksell, 1972.

Merricks, A.P. and Crocker, R.K. The Influence of Science Process Activities and Selected Science Reading Materials on Reading Achievement of First and Third Grade Pupils. School Science and Mathematics (In Press).

Mayfield, J.M. Factors affecting rationality in the discussion of a problem by small groups of secondary school students. Science Education, 60, 173-184, 1976.

Osgood, C.E. and Suci, G.J. A Measure of relation determined by both mean differences and profile information. Psych. Bull., 49, 251-262, 1962.

O'Leary, K.D. and O'Leary, S.G. Classroom Management. New York: Pergonon Press, 1972.

Piaget, J. The Psychology of Intelligence. Totowa, New Jersey: Littlefield Adams, 1973.

Porter, R.B. and Cattell, R.B. Handbook for the Children's Personality Questionnaire. Institute for Personality and Ability Testing, Illinois, 1968.

Power, C.N. The Unintentional Consequences of Science Teaching. Journal of Research in Science Teaching, 10, 331-339, 1973.

Programme on Teaching Effectiveness. A Factorially Designed Experiment on Teacher Structuring, Soliciting, and Reacting. Stanford University, Stanford California: Stanford Centre for Research and Development in Teaching. Research and Development Memorandum, No. 147, November, 1976.

Raven, J.C. Guide to the Standard Progressive Matrices Sets A, B, C, D and E. H.K. Lewis & Co. Ltd., London, 1960.

Rosenshine, B. Classroom Instruction. In N.L. Gage (Ed.). The Psychology of Teaching Methods. 77th yearbook of the National Society for the Study of Education, 1976.

Soar, R.S. and Soar, R.M. Research Reports. Classroom Behavior, Pupil Characteristics and Pupil Growth for the School Year and the Summer. University of Florida, December, 1973.

Soar, R. Final Report, Follow though classroom process measurement and pupil growth (1970-1971). Gainsville, Fla.: Institute for the Development of Human Resources, University of Florida, 1973.

Solomon, D. and Kendall, A.J. Final Report. Individual Characteristics and Children's Performance in Varied Educational Settings. Psychological Services Section, Montgomery County Public Schools, Rockville, Maryland, May, 1976.

Stallings, J. Implementation and Child Effects of Teaching Practices in Follow Through Classrooms. Monographs of the Society for Research in Child Development. Serial No. 163, Vol. 40, No. 7-8, December, 1975.

Shymansky, J.A. and Matthews, C.C. A Comparative Laboratory Study of the Effects of Two Teaching Patterns on Certain Aspects of the Behavior of Students in Fifth Grade Science. *Journal of Research in Science Teaching*, Vol. 11, No. 2, 157-168, 1974.

Schwab, J. "The Structure of the Natural Sciences" In Ford and Pugno (Eds.). *The Structure of Knowledge and the Curriculum*. Chicago: Rand McNally, 1964.

Schutz, W.C. *FIRO: A Three Dimensional Study of Interpersonal Behavior*. New York: Rinehart, 1958.

Stallings, J. and Kaskowitz, D. Follow through Classroom Observation evaluation, 1972-1973. *Menlo Park, Cat.: Stanford Research Institute, 1974*. Office of Education Contract OEC-0-8522480-4633(100).

Snow, R. Representative and quasi-representative designs for research on teaching. *Review of Educational Research*, 44, 265-291, 1974.

Tannenbaum, R.S. The Development of the Test of Science Process. *Journal of Research in Science Teaching*, Vol. 8, 123-136, 1971.

Tobias, S. Achievement treatment interactions. *Review of Educational Research*, 46, 61-74, 1976.

Wittgenstein, L. *Philosophical Investigations*. Oxford, Basil Blackwell, 1958.